

DEPARTMENT OF COMPUTER SCIENCE
SERIES OF PUBLICATIONS A
REPORT A-2018-4

Exploring the Dynamics of the Biocybernetic Loop in Physiological Computing

Ilkka Kosunen

*To be presented, with the permission of the Faculty of Science of
the University of Helsinki, for public examination in room D122,
Exactum, on 22th of March 2018, at 12 o'clock noon.*

UNIVERSITY OF HELSINKI
FINLAND

Supervisor

Giulio Jacucci, University of Helsinki, Finland

Pre-examiners

Thorsten O. Zander, Technische Universität Berlin, Germany

Erin Solovey, Drexel University, USA

Opponent

Stephen Fairclough, Liverpool John Moores University, United Kingdom

Custos

Giulio Jacucci, University of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Gustaf Hållströmin katu 2b)
FI-00014 University of Helsinki
Finland

Email address: info@cs.helsinki.fi

URL: <http://cs.helsinki.fi/>

Telephone: +358 2941 911, telefax: +358 9 876 4314

Copyright © 2018 Ilkka Kosunen

ISSN 1238-8645

ISBN 978-951-51-4138-5 (paperback)

ISBN 978-951-51-4139-2 (PDF)

Computing Reviews (1998) Classification: H.5.2

Helsinki 2018

Unigrafia

Exploring the Dynamics of the Biocybernetic Loop in Physiological Computing

Ilkka Kosunen

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
ilkka.kosunen@gmail.com
<http://cs.helsinki.fi/>

PhD Thesis, Series of Publications A, Report A-2018-4
Helsinki, March 2018, 91 + 161 pages
ISSN 1238-8645
ISBN 978-951-51-4138-5 (paperback)
ISBN 978-951-51-4139-2 (PDF)

Abstract

Physiological computing is a highly multidisciplinary emerging field in which the spread of results across several application areas and disciplines creates a challenge of combining the lessons learned from various studies. The thesis comprises diverse publications that together create a privileged position for contributing to a common understanding of the roles and uses of physiological computing systems, generalizability of results across application areas, the theoretical grounding of the field (as with the various ways the psychophysiological states of the user can be modeled), and the emerging data analysis approaches from the domain of machine learning.

The core of physiological computing systems has been built around the concept of *biocybernetic loop*, aimed at providing real-time adaptation to the cognitions, motivations, and emotions of the user. However, the traditional concept of the biocybernetic loop has been both self-regulatory and immediate; that is, the system adapts to the user immediately. The thesis presents an argument that this is too narrow a view of physiological computing, and it explores scenarios wherein the physiological signals are used not only to adapt to the user but to aid system developers in designing better systems, as well as to aid other users of the system.

The thesis includes eight case studies designed to answer three research questions: 1) what are the various dynamics the biocybernetic loop can

display, 2) how do the changes in loop dynamics affect the way the user is represented and modeled, and 3) how do the choices of loop dynamics and user representations affect the selection of machine learning methods and approaches? To answer these questions, an analytical model for physiological computing is presented that divides each of the physiological computing systems into five separate layers.

The thesis presents three main findings corresponding to the three research questions: Firstly, the case studies show that physiological computing extends beyond the simple real-time self-regulatory loop. Secondly, the selected user representations seem to correlate with the type of loop dynamics. Finally, the case studies show that the machine learning approaches are implemented at the level of feature generation and are used when the loop diverges from the traditional real-time and self-regulatory dynamics into systems where the adaptation happens in the future.

Computing Reviews (1998) Categories and Subject Descriptors:

H.5.2 User Interfaces

General Terms:

Physiological Computing, Human-Computer Interaction

Additional Key Words and Phrases:

HCI

Acknowledgements

I would like to thank Professor Giulio Jacucci for supervising this work and providing me both guidance when needed as well as academic freedom to pursue my own ideas.

My research received financial support from the European Commission through the Mindsee and CEEDs projects as well as from the Finnish Funding Agency for Technology and Innovation through the Emokeitai project. I also owe thanks to Department of Computer Science at the University of Helsinki, as well as Helsinki Institute for Information Technology (HIIT) for providing the infrastructure and support that enabled me to conduct this research.

I would like to give special thanks to Matti Luukkainen, Petri Lievonen, Petri Savolainen and Matti Nelimarkka who gave invaluable feedback on the thesis. I'm especially grateful for Antti Salovaara who tirelessly answered all my questions. I would like to give special thank to Pirjo Moen for helping me with all the practicalities of my doctoral studies.

I would also thank all my colleagues at Department of Computer Science and the Ubiquitous Interaction group: Antti Jylhä, Imtiaz Ahmed, Khalil Klouche, Baris Serim, Tuukka Ruotsalo, Kumaripaba Athukorala, Chen He, Yi-Ta Hsieh, Andrea Vieanello, Tung Vuong. I would like to thank the CKIR group for guiding me in the use of psychophysiological recordings: Niklas Ravaja, Mikko Salminen, Simo Järvelä and Matias Kivikangas.

I'm extremely grateful for all my co-authors without whom this work would not have been possible: Kai Kuikkaniemi, Benjamin Cowley, Manuel J. A. Eugster, Michiel M. Spapé, Samuel Kaski, Petri Lankoski, Inger Ekman, Jaakko Kemppainen, Antti Ruonala and Tetsuo Yamabe. I'm especially grateful for two of my co-authors whom with I collaborated extensively: Jussi Palomäki and Oswald Barral.

I would also like to thank all my colleagues at Helsinki Institute for Information Technology that I've had the pleasure to work and interact with: Airi Lampinen, Antti Oulasvirta, Antti Ukkonen, Asko Lehmuskallio, Ella Bingham, Eve Hoggan, Dorota Glowacka, Hannu Toivonen, Herkko Hi-

etanen, Joanna Bergström-Lehtovirta, Joel Pyykkö, Jouni A. Ojala, Juho Hamari, Kai Huotari, Lassi Liikkanen, Petri Myllymäki, Marco Filetti, Martti Mäntylä, Max Vilkki, Mikael Johnson, Olli Pitkänen, Patrik Florén, Sasu Tarkoma, Tapio Takala, Teemu Roos, Tuukka Lehtiniemi, Yves Florack, Vesa Kantola, Vili Lehdonvirta and Vilma Lehtinen.

Very "special" thanks go to the cross-scientific art-science research colloquim Mäyränkaatajat for supporting me also in my personal life when I needed break from the research activities. I'm also grateful for my parents and siblings for their continuous support.

Helsinki, March 2018
Ilkka Kosunen

Contents

1	Introduction	1
1.1	Scope of the Thesis	4
1.2	Objectives: Exploring Physiological Computing from Three Perspectives	6
1.3	Structure of the Thesis	7
2	Background	9
2.1	Psychophysiology	9
2.2	Physiological Signals	10
2.2.1	Electrodermal Activity	10
2.2.2	Electromyography	12
2.2.3	Electrocardiography	12
2.2.4	Electroencephalography	14
2.3	Physiological Computing	15
2.3.1	The Four Categories of Physiological Computing	16
2.4	Affective Computing	17
2.4.1	Emotion Theories and Definitions	18
2.4.2	Decision-Making and Emotions	19
2.5	Wearable Computing	21
3	Research Questions and Method	23
3.1	Research Methods	23
3.1.1	Design Science	24
3.1.2	Research Design: Case Studies	25
3.2	Research Questions	27
3.3	The Five-Layer Model of Physiological Computing	30
3.3.1	The Signal Layer	31
3.3.2	The Metrics Layer	31
3.3.3	The Indices Layer	32
3.3.4	The Logic Layer	33
3.3.5	The Application Layer	33

3.3.6	The Full Model	33
4	Empirical Studies	37
4.1	Games	37
4.1.1	Background and Related Work	38
4.1.2	The Studies	40
4.1.3	Summary	46
4.2	Physiological Annotation and Information Retrieval	46
4.2.1	Background and Related Work	46
4.2.2	The Studies	49
4.2.3	Summary	60
4.3	Meditation	61
4.3.1	Background and Related Work	61
4.3.2	The Study	63
4.3.3	Summary	65
5	Findings	67
5.1	Purpose: Extending the Concept of the Biocybernetic Loop	67
5.2	Representation: How User State Is Modeled	69
5.3	Approaches: When and How to Use Machine Learning . . .	71
6	Discussion	73
6.1	Summary of the Main Findings	74
6.1.1	Implications of the Research	75
6.1.2	Limitations	76
6.1.3	Directions for the Future	76
	References	79

List of Publications

The thesis consists of a summarizing overview and the following publications, referred to as publications I–VII in the text. These publications are reproduced at the end of the thesis.

- I. **Ilkka Kosunen**, Mikko Salminen, Simo Järvelä, Antti Ruonala, Niklas Ravaja, Giulio Jacucci. RelaWorld: Neuroadaptive and immersive virtual reality meditation system. *Proceedings of the 21st International Conference on Intelligent User Interfaces IUI 2016*.

Contribution: The design of a virtual-reality meditation system was jointly planned with the RelaWorld Team. During the development of the system, the author was responsible for real-time analysis and handling of the EEG signals. The author analyzed the data jointly with Mikko Salminen and wrote the first draft of the paper. All authors contributed to the revision of the paper.

- II. Jussi Palomäki, **Ilkka Kosunen**, Kai Kuikkaniemi, Tetsuo Yamabe, Niklas Ravaja. Anticipatory electrodermal activity and decision making in a computer poker-game. *Journal of Neuroscience, Psychology, and Economics* 6 (1), 55, 2013.

Contribution: The author was responsible for the recording and analysis of the physiological signals (EDA and EKG), as well as helping in the modification of the poker game software to fit the user study requirements. The author was responsible for the data analysis and contributed to the writing of the first draft of the paper. All authors contributed to the revision of the paper.

- III. Manuel J. A. Eugster, Tuukka Ruotsalo, Michiel M. Spapé, **Ilkka Kosunen**, Oswald Barral, Niklas Ravaja, Giulio Jacucci, Samuel Kaski. Predicting term-relevance from brain signals. *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 425–43, 2014.

Contribution: The author took part in designing the experiment and in the data analysis. All authors contributed to the revision of the paper.

- IV. Oswald Barral, **Ilkka Kosunen**, Tuukka Ruotsalo, Michiel M. Spapé, Manuel J. Eugster, Niklas Ravaja, Samuel Kaski, Giulio Jacucci. Extracting relevance and affect information from physiological text an-

notation. *User Modeling and User-Adapted Interaction* 26, no. 5 (2016): 493–520.

Contribution: The author took part in designing the experiments, including creating a web-proxy sever to collect the physiological data and synchronize said data with the behavioral responses. The author also participated in the data analysis and in drafting of the first version of the paper. All authors participated in the revision.

- V. Oswald Barral*, **Ilkka Kosunen***, Giulio Jacucci. No need to laugh out loud: Predicting humor appraisal of comic strips based on physiological signals in a realistic environment, *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(6), 40. .

Contribution: The author designed the experiment together with Oswald Barral, as well as participated in the data-analysis and writing of the first draft of the paper. All authors contributed to the revision of the paper. *The first two authors shared equal contribution.

- VI. Benjamin Cowley, **Ilkka Kosunen**, Petri Lankoski, J. Matias Kivikangas, Simo Järvelä, Inger Ekman, Jaakko Kemppainen, Niklas Ravaja. Experience assessment and design in the analysis of gameplay. *Simulation and Gaming*, 45 no. 1: 41–69, February 2014.

Contribution: The author was responsible for the unsupervised clustering of the physiological data as well as the game events, also participating in the writing of the paper.

- VII. **Ilkka Kosunen**, Jussi Palomäki, Giulio Jacucci, Niklas Ravaja. Heart-rate sonification biofeedback for poker. *Submitted to International Journal of Human-Computer Studies*.

Contribution: The author took part in designing the experiment and in modifying the poker game system to support audio biofeedback and facilitate the user experiment. The author also performed the data analysis and wrote the first draft of the paper with Jussi Palomäki. All authors contributed to the revision of the paper.

Chapter 1

Introduction

Physiological computing is a form of HCI wherein the interaction depends on measuring and responding to the physiological activity of the user in real time (Fairclough, 2009). Physiological computing has the potential to revolutionize human–computer interaction. For one thing, it increases the communication bandwidth dramatically by introducing several new information channels. It also enables the computer to sense the implicit and affective signals of the user, thereby creating possibilities for technologies such as affective computing (Picard, 1997). The traditional communication between humans and computers has been described as asymmetrical (Hettinger et al., 2003): while the computer is able to output vast amounts of audiovisual information quickly, the input from the user is limited to the relatively low bandwidth provided by mouse and keyboard. Furthermore, while the user has access to the internal state of the computer system (e.g., memory consumption and processor utilization levels), the computer has no information on the cognitive and emotional state of the user (Fairclough, 2009). Physiological computing allows symmetry in terms of both information bandwidth (the added input modalities in the form of physiological signals) and the user state and context information derived from the physiological data.

The standard mode of human–computer interaction has been completely explicit: the computer reacts only to explicit commands given to it by the user. Physiological computing also enables *implicit* communication; by observing the physiological signals of the user, the computer can detect, for example, when the task the user is performing is too challenging and automatically decrease the difficulty level, or when users are getting distracted from the task, the system could give them a notification. Successful examples of applying the latter principle include detection of driver fatigue (Lal and Craig, 2001) and mental workload of operators (Boyer et al., 2015).

Measuring the implicit feedback from the user has implications for several application areas. User interface designers can automatically evaluate their designs: mental workload of drivers (Solovey et al., 2014) can be measured in real time to evaluate the cognitive demands of different interface designs. Implicit signals can also be used for automated annotation of data: the computer can automatically detect whether a document contains relevant information (Eugster et al., 2014; Barral et al., 2015; Ruotsalo et al., 2014) and what kind of affective reaction the content generated in the user.

Physiological computing can also benefit the development of interactive systems. One of the dominant themes in design is human-centered design, which is concerned with incorporating the user’s perspective into the software development process in order to achieve a usable system (Maguire, 2001). However, the HCD approach has been criticized as potentially harmful for trying to optimize the design for a generic “everyman” user. What is needed is the *individuation* (Hancock et al., 2005) of the interface to the needs of each specific user, something that can be accomplished by measuring the implicit feedback from the user: since the system can adapt in real time to the user’s implicit reactions, the system can continuously calibrate itself to provide the optimal user experience for each individual user.

Measuring the psychological state of the user creates intriguing possibilities for computer games (Spapé et al., 2015b, 2013). Games can automatically adjust their difficulty to the best possible fit for the individual user’s skill level. The game can also detect affective phenomena such as when the user is afraid and use this information to create impactful events or perhaps adjust the music and background sounds. Even the narrative can be tied to the physiological responses: the game might detect how the player responds to different characters in the game world and choose to have certain events affect the characters that generated the strongest affective reactions in the player. The physiological signals can also be directly tied to the game mechanics (Nacke et al., 2011): the game can be designed such that the player character’s speed increases as the player gets aroused, or accuracy and aim ability might increase when the player calms down (Kuikkaniemi et al., 2010). Apart from games, physiological computing has been utilized in interactive art (Edmonds et al., 2004).

However, the great potential of physiological computing can lead to unsubstantiated optimism: topics such as brain–computer interfaces can lead to fantastical claims of computers capable of mind-reading that have little scientific validity (Spapé et al., 2015a). A rigorous scientific approach that takes into account the potential issues with reliability and validity of physiological recordings is necessary. Furthermore, the field of physiological

computing is still highly unorganized: each new project is usually started from scratch, and there is little reuse of code, resources, and best practices. Also, the research papers are poor design documents or programming specifications: it is often practically impossible to replicate the experimental setup of another research team. Properly organized and structured documentation along with well-specified ways of sharing resources such as code and data could reduce the development time and provide a way of comparing results between research groups.

This thesis was written to respond to the aforementioned challenges by examining a broad set of physiological computing systems from three distinct perspectives: Firstly, I explore the different roles that physiological computing can take, which range from self-regulation to facilitating technology design. Secondly, I explore the various ways users are modeled in physiological computing. Thirdly, I compare the approaches taken to building physiological computing systems, which range from hand-tuned “expert systems” to various machine learning approaches. After this, the differences and similarities related to these topics across applications are discussed. To facilitate this discussion, a layered analytical model is developed to explicate the individual aspects of physiological computing.

The thesis examines how physiological computing can be applied in real-life applications that range from the work-related and more serious to games and play, as well as to health and wellness; they thus cover the full spectrum of what might be considered the needs and situations of an average person. In the thesis, we examine what kinds of applications are available for each of these fields, as well as which physiological signals are most suitable in each case, also considering whether there are case-specific differences in how these signals are best interpreted. Instead of being just a review, the thesis includes full examples, with detailed user studies, for each of the application areas discussed. We then explore how these various application areas can be seen through the lens of the three perspectives outlined above and what similarities and differences arise on the basis of the application context.

Physiological computing is closely related to fields such as ubiquitous computing (Abowd et al., 2002), pervasive computing (Satyanarayanan, 2001), ambient intelligence (Ramos et al., 2008), enactive interaction (Kaipainen et al., 2011), affective computing (Picard, 1997), and symbiotic interaction (Jacucci et al., 2014). This thesis also contributes to these fields, and we hope that it will show how work on these topics can be mutually supportive across disciplines.

1.1 Scope of the Thesis

Physiological computing is a complex topic that cannot be exhaustively covered in a single thesis. Therefore, to keep the task feasible, the scope of the thesis is constrained in two ways. Firstly, it is limited on a practical level: only technologies that could potentially be used by normal users in their everyday life outside the laboratory are considered; clinical uses of physiological computing, such as brain–computer interfaces for locked-in patients and myoelectric-controlled human arm prosthetics, are not covered. Furthermore, only non-invasive sensors that could potentially be used in real-life situations are dealt with. This rules out technologies such as sensors that are implanted directly in the brain (intracranial EEG) and sensors that for other reasons are unsuitable for practical applications. The latter category includes most brain imaging techniques, such as fMRI.

Secondly, the scope is limited on a theoretical level to cover only parts of the field of physiological computing. In his seminal paper, Fairclough (2009) identified six fundamental issues for physiological computing. For the positioning of this thesis within the field of physiological computing, these fundamental issues are briefly described, and then a description is given of how we attempt to address each of the issues.

Psychophysiological Inference

There rarely exists a one-to-one mapping between a physiological signal and a psychological state of the user (Cacioppo et al., 2007). That is, each physiological signal usually can be an indicator of several physiological states, and each state can be inferred from multiple sources.

Psychophysiological inference is a complex topic that cannot be comprehensively covered in this thesis, but it is an important field of basic research in psychophysiology that this thesis will build upon. While we do not aim to address it directly, the thesis does provide a small contribution in relation to this issue by describing how the physiological inference was implemented in the publications included in the thesis.

Psychophysiological Validity

Once the setup for the physiological inference has been decided on, it needs to be validated. Properly validating a given inference pattern is a basic research matter demanding careful and systematic empirical studies, and this is not the aim for the thesis. While it is not the main goal, several types of psychophysiological inferences are validated in the publications – for

example, how in a scientific search context certain brain activity patterns indicate relevance.

Representation of the User

For the psychophysiological inferences to be useful for application design, they must be *operationalized* in a way that allows applications to utilize them. There exist multiple ways to represent the user state, but the approaches can be divided into two main groups: in dimensional models the user state is represented as a point in a space spanned by some basic indexes such as arousal and valence, while in categorical representations the user is classified as being in a specific state (e.g., being angry or happy). It is also possible to utilize machine learning methods that automatically generate (often a black-box) representation of a user. The publications presented in this thesis cover both of these approaches and show how they can be successfully used in real-life applications.

Awareness and Interaction Design

The next fundamental issue has to do with the types of adaptations the system can perform. In the original formulation by Fairclough (2009), the question was centered on implicit vs. explicit interaction: should the physiological computing system give explicit feedback to the user or adapt in an implicit manner instead? However, the discussion was always of self-regulation: the system adapted to the physiological state of the user. In this thesis, the question is expanded to pertain to not only self-regulation but also situations wherein the physiological signals of the user are used to not only enhance the current user's experience but also aid designers of the system, as in the case of technology design, and other users too, as with automated content annotation.

Dynamics of the Biocybernetic Loop

The biocybernetic loop forms the core of the physiological computing system. As defined by Fairclough, the functional goal of the loop is to “derive real-time adaptations to cognitions, motivations, and emotions that appear both timely and intuitive from the users' perspective” (Fairclough, 2009). However, with this thesis we aim to expand the concept of the loop in two ways: instead of directing the feedback loop directly to the user, we examine physiological computing systems in which the feedback goes to, firstly, the designers of the system or, secondly, other users of the system.

The main contribution of this thesis is in exploring several ways the biocybernetic loop can be constructed. Special attention is paid to the question of when the loop should be designed by hand as an “expert system” and when machine learning should be used to find optimal ways to utilize specific physiological signals.

Ethical Implications

Ethics considerations are an important part of physiological computing but are not within the scope of this thesis.

1.2 Objectives: Exploring Physiological Computing from Three Perspectives

As noted above, the objective for the thesis is to survey the field of physiological computing from three perspectives: the purpose or role of physiological computing, the representation of the user, and the various approaches to the design of physiological computing systems. These three perspectives are tied closely to the three research questions for this work, which are described in detail in Chapter 3.

Purpose:

The first perspective is that of the different roles of physiological computing. Traditionally, physiological computing has been used for self-regulating: the system adapts to the user in line with the physiological signals of that user. However, in this thesis we propose that physiological computing encompasses much more. For Publication VI, physiological computing was used for technology design via clustering physiological signals of players to recognize interesting behavioral patterns. For publications IV and V, physiological signals were used to annotate content. Such annotations help not only the user; they can be used also to train recommender systems to help other users, or even aid the computer in developing a sense of humor as described in Publication V. With the first research question, described alongside the other two in Section 3.2, we tried to provide insights into this phenomenon by exploring the different forms the underlying biocybernetic loop can take.

Representation:

The second perspective deals with the different ways users are modeled and represented in physiological computing applications. In biofeedback, which could be considered the simplest form of physiological computing, the physiological signals are directly mapped to audiovisual output. Often,

though, the physiological signals are first interpreted and operationalized as *indices* of cognitive and affective states that are then used as the input to the adaptive system. These states can range from the simple interpretation of sympathetic nervous system activity such as arousal to complex mappings of brain activity to cognitive states, possibly involving several stages of interpretation and fusion of numerous physiological signals, sometimes with other context information. The second objective set for this thesis was to explore the various representations and their uses in a wide range of applications, to discover the similarities and differences. Accordingly, the second research question deals with how the representations are altered as the dynamics of the biocybernetic loop change.

Approaches:

The third perspective is related to the internal logic of the physiological computing applications. In the simplest case of biofeedback, there exists a pre-defined and clear mapping of a signal to output, such as giving audio feedback when a certain signal rises above a threshold value. Often, a more complicated decision mechanism is needed, such as classifying affective state from a combination of several physiological sources. In this case, it might be useful to apply machine learning techniques to determine what exactly the rules are according to which certain signals should indicate some affective states, such as “joy” or “anger.” In the most extreme case, we might just give the machine learning algorithm an unlabeled set of physiological data and ask the machine to find potentially interesting behavioral patterns. The publications cover all of these approaches, with the third objective for the thesis being to survey how these different approaches are used, what their benefits are, and how they differ. The third research question reflects this perspective by directing us to examine how the approaches change when the dynamics of the biocybernetic loop vary.

To facilitate the three-perspective view of physiological computing, an analytical framework is presented that simultaneously displays the three distinct facets and presents the similarities and differences between the included publications from all three vantage points. The analytical framework is constructed in a modular fashion with the aim of explicating those commonalities between applications that could be easily reused and those that are application-specific.

1.3 Structure of the Thesis

Continuing the groundwork laid in the introductory chapter, Chapter 2 provides an overview of physiological computing and related fields such

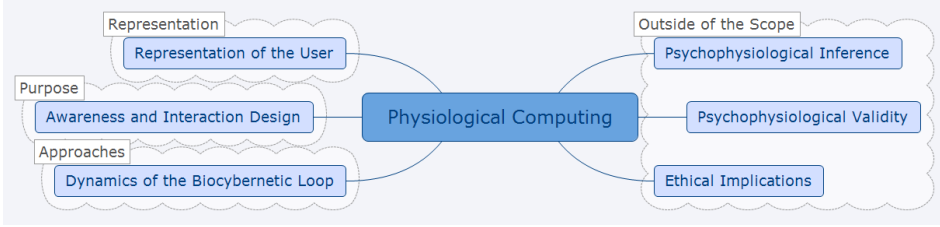


Figure 1.1: The scope of the thesis, covering three of the six fundamental issues.

as affective computing, while also addressing the underlying theoretical grounding of psychophysiological research and giving a brief description of various physiological signals. In Chapter 3, the research methods and research questions are presented. Then, Chapter 4 presents an overview of the research reported on in publications I–VII. Findings and answers to the research questions are presented in Chapter 5, and the final chapter rounds out the work with conclusions and discussion considering such matters as limitations and implications for future research.

Chapter 2

Background

Physiological computing is a multidisciplinary and variety-rich field. The background presented first describes the theoretical grounding of the field in extensive basic research in the field of psychophysiology. Then a brief description is offered of the most important physiological signals, along with how they are used in physiological computing, because at least a rudimentary grasp of the physiological signals and their analysis is necessary for understanding the rest of the thesis. A survey of related fields such as affective computing and wearable computing is given to position the work in the thesis within the larger framework of related technologies. Then we discuss the new machine learning approaches that are changing the approach to development of physiological computing systems.

2.1 Psychophysiology

Physiological computing is largely based on work done in the field of psychophysiology. John L. Andreassi gives the following definition of psychophysiology: “Psychophysiology is the study of relations between psychological manipulations and resulting physiological responses, measured in the living organism, to promote understanding of the relation between mental and bodily processes” (Andreassi, 2000). Another way to put it would be to say that psychophysiology is a combination of anatomy, physiology, and psychology, and very closely related to behavioral neuroscience (Cacioppo et al., 2007).

What is especially relevant for physiological computing is the way psychophysiology has been used in media research, since this constitutes, in effect, half of the equation: the user’s physiological response to multimedia content. The work in media research has taken into account the fact that

a multimodal stimulus, which is usually involved when one is working with computers, can be more complicated to interpret than, say, single-tone beeps as might be used in more traditional psychophysiological research (Ravaja, 2004).

2.2 Physiological Signals

In this thesis, we concentrate on the four most common physiological signals and their analysis and interpretation. While the thesis is not about signal analysis, these are such a central part of physiological computing that without a basic understanding of how they operate, it would be difficult to appreciate the more specialized application of physiological computing that is demonstrated in the publications.

2.2.1 Electrodermal Activity

Electrodermal activity (EDA), also known as skin conductance (SC), is, to put it simply, measurement of sweating or, to be more precise, a measure of the changes in electrical properties of the skin due to sweating. Human skin contains two types of sweat glands: the apocrine and the eccrine. The apocrine glands, which are found in the genital areas and armpits, have to do with thermal regulation and are not of interest in a psychophysiological context. The eccrine glands, located in the palms of the hands and soles of the feet, are of greater interest, because their activation is connected more closely with emotional reactions than with temperature (Andreassi, 2000).

Traditionally, EDA has been seen as an indicator of arousal: arousal causes the activation of eccrine glands, which increases skin conductivity – that is, EDA (Bradley and Lang, 2000). However, EDA is also linked to specific short-term events, such as an orienting reaction to novel stimuli, mental workload, and cognitive appraisal of a stimulus. For Publication IV, EDA was used to classify content as “relevant.” This signal can be useful for a large number of use cases, but, since it is connected to so many psychological events, care must be taken not to mistakenly interpret an orienting response as arousal, for example.

The EDA signal is a combination of two components: a slow, underlying tonic signal and a faster phasic component that consists of event-related “spikes” in the signal (Benedek and Kaernbach, 2010). In simple terms, the larger the phasic spike in the EDA signal, the larger or stronger the stimulus that caused it. Accordingly, analysis of the EDA signal usually involves first trying to remove the effect of the underlying tonic component, which is a combination of the underlying mood and stress level of the user,

or perhaps changing external conditions, such as the temperature. There are two main ways of doing the extraction. One is a crude method wherein each spike is centered at its mean. There are some problems with the simple approach, mainly that of overlapping phasic spikes. A single phasic spike has a shape that slopes upward rapidly and then slowly returns to the tonic level, but if the user experiences several stimuli or events in fast progression, there can be several phasic spikes, each building on top of the one before it. Deconvolution-based algorithms exist that can separate the phasic spikes from one another, but they are somewhat slow and are not really suitable for a real-world application that needs to process the spikes as they occur (Benedek and Kaernbach, 2010).

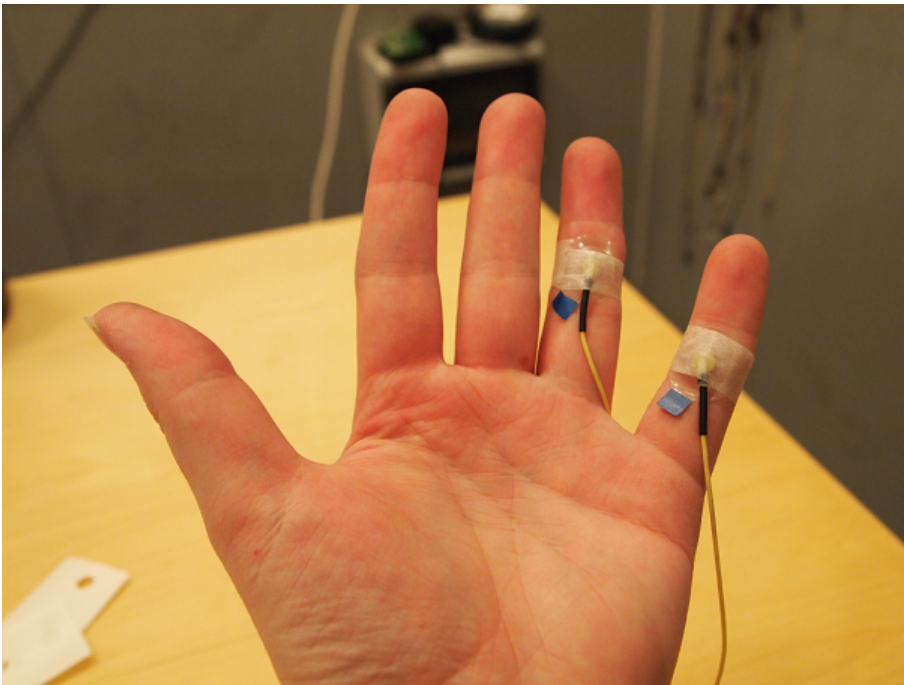


Figure 2.1: EDA sensor attached to the medial phalanges of the ring and little finger.

In the studies reported on in the publications, EDA was used to predict investment decisions in a poker game (Publication II); as one of the signals used to cluster user behavior for game design (Publication VI); and to implicitly annotate relevance (Publication IV), affects (Publication IV), and humor (Publication V).

2.2.2 Electromyography

Electromyography refers to measuring and recording of muscle potentials, specifically the activity associated with muscle contractions (Tassinari and Cacioppo, 2000). The recording can be done either by placing an electrode needle directly into the muscle or via a surface electrode. In HCI settings, the surface method is almost always chosen, on account of its noninvasive nature. The EMG signal is generated by muscle action potentials spreading over skeletal cells after a neural stimulation. Detection of a momentary difference in potential between electrodes spaced over the muscle indicates a wave of depolarization following a muscle contraction. In a skeletal muscle, all cells fire simultaneously when the muscle is activated, yet the distance from cells in each part of the muscle to the electrodes varies, thereby causing the EMG signal to be not a single spike but a wave of signals arriving slightly apart. Accordingly, the signal is usually taken as an integrated value over time.

While any muscle could provide interesting opportunities for physiological computing, much of the research has traditionally concentrated on facial muscles, since these give fast, reliable, and accurate indicators. Indeed, facial EMG can even detect muscle activity that is not visually perceptible (Ravaja, 2004). Three specific muscles have become the *de facto* locations in psychophysiological research: the corrugator supercilii (CS), above the eye; the orbicularis oculi (OO), beneath the eye; and the zygomaticus major (ZM), in the cheek. Activity of the CS has been linked to negative valence and concentration, while that of the other two indicates positive valence.

EMG was used for Publication IV’s study, in which CS activity was used to annotate relevance of articles, and for Publication VI, for which it, CS, OO, and ZM were used to cluster user behavior for game design.

2.2.3 Electrocardiography

Heart rate is arguably the most well-known of the physiological metrics, and the relationship between heart activity and emotions has been known since ancient times. In psychophysiology, the research usually centers on the interpretation of heart rate and its aggregates, such as heart-rate variability. There are several possible ways of recording heart rate, among them photoplethysmography (PLG) (Allen, 2007), in which the recording is done by passing light through tissue such as a finger, and ballistocardiography, which can measure heart rate from the chair or bed a user is resting on by detecting the minute mechanical movement of the user’s body caused by

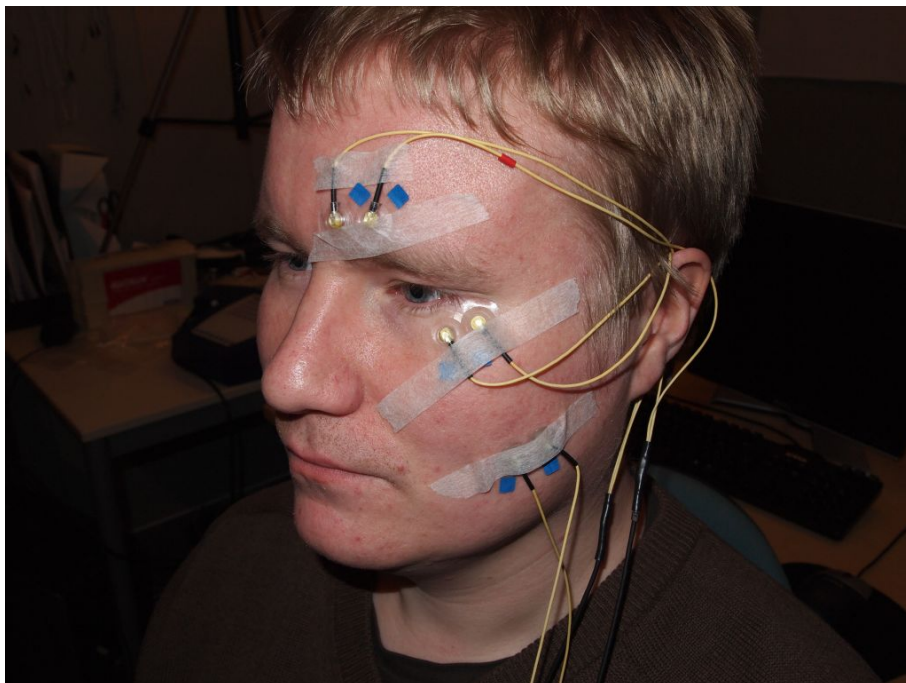


Figure 2.2: Facial EMG sensors for, from top to bottom, the corrugator supercilii, orbicularis oculi, and zygomaticus major muscles.

each heartbeat (Anttonen and Surakka, 2005). Recently it has been shown that heart rate can even be detected with a low-cost web cam and facial video recordings through detection of subtle light changes in a fashion similar to PLG's (Bousefsaf et al., 2014). However, by far the most common technique for measuring heart rate is to record the electrical activity of the heart by using electrocardiography (EKG). From a practical perspective, heart rate is a highly useful signal because it is already recorded by a wide variety of consumer fitness devices (Gamelin et al., 2006).

The heart is unique in that it is influenced by both the sympathetic and the parasympathetic nervous system. The sympathetic component, which deals with “fight-or-flight” responses, usually increases the heart rate, while the parasympathetic system, responsible for “rest-and-digest” behavior, tends to decrease it (Cowley et al., 2016). Common features extracted from ECG are heart rate (HR) and heart-rate variability (HRV) – which can indicate, for example, mental workload (Cowley et al., 2016).

In Publication VII, ECG is used to create sonified heart-rate audio biofeedback to assist poker players.

2.2.4 Electroencephalography

Recording the brain activity directly has obvious appeal for applications that depend on some mental or cognitive index such as perceived relevance, as compared with signals of peripheral physiology (e.g., EDA), which are only indirect expressions of the cognitive/affective reactions that occur in the brain. There are several ways of measuring the brain activity directly. Among them are magnetic resonance imaging (MRI), magnetoencephalography (MEG), positron emission tomography (PET), functional near-infrared spectroscopy (fNIRS), and electroencephalography (EEG). Some of the imaging technologies, such as MRI, MEG, and PET, require bulky equipment and are not suitable for physiological computing. While fNIRS has been successfully applied in fields such as brain–computer interfaces and shows a large amount of potential (Solovey et al., 2015), most research and development in physiological computing surrounds EEG, and indeed EEG was the technology used for the publications.

Analysis of the EEG signal is divided sharply into two kinds: the signal can be analyzed in either the frequency or the time domain. In frequency-domain analysis, rhythmic oscillations in the brain are measured. The exact cause of these oscillations is still debated, but they have been found to be relevant for at least coding information, setting and modulating brain attentional states, and ensuring communication between neuronal populations around the brain (da Silva, 2013). Several specific oscillation bands have been defined, such as the delta (0.2–3.5 Hz), theta (4–7.5 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (30–90 Hz), and each has its own interpretations for metrics. For example, increased alpha activity is an indicator of relaxation. Also, the difference in oscillation between certain sites in the brain can be an important metric, as in the case of frontal asymmetry. The domain of frequency-based EEG analysis is vast, and the interested reader is directed to other work (Cowley et al., 2016; Cacioppo et al., 2007; da Silva, 2013) for details. For Publication I, alpha and delta activity were used as biofeedback signals, while for Publication V, gamma-band activity was found to be highly correlated with how funny people found media content they were browsing.

The other way to interpret the EEG signal is to analyze it in the time domain, ignoring any frequency patterns. These time-locked analyses deal with what are known as event-related potentials (ERPs), which are neuronal potentials that occur in a certain time window after an event has occurred. Usually, these potentials are so weak – and the signal so full of noise – that a stimulus is presented multiple times and the EEG response averaged to get a statistical ERP mean. All ERPs are named on the basis

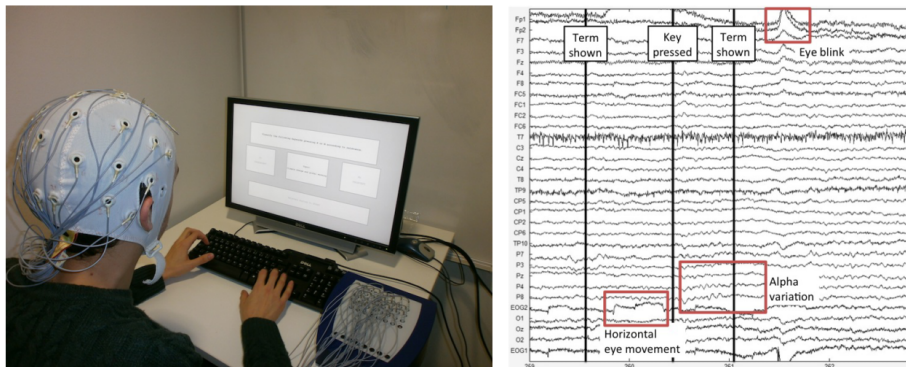


Figure 2.3: Example of EEG recording from Publication III.

of whether the potential is negative or instead positive and with an indication of the associated time delay, so, for instance, a positive ERP that occurs 300 milliseconds after the event is labeled as P300. A very large amount of research has been done in relation to the various ERPs, something that Luck refers to in his book as “ERPology” (Luck, 2014). Detailed description of these too is beyond the scope of this thesis.

We used ERP-based analysis for Publication III, where it was utilized to predict the relevance of keywords.

2.3 Physiological Computing

Physiological computing is a mode of HCI wherein the interaction depends on measurement of and response to the physiological activity of the user in real time (Fairclough, 2009). This process usually takes place as a series of steps: Firstly, one of the signals described earlier in the chapter is selected – for instance, skin conductance. Then some quantifiable *metrics* or *features* are calculated from the signal, such as the amplitude of a phasic spike in the case of EDA. These metrics then are usually interpreted as representing some user state (such as arousal), and some logic for how the system should react to that state is decided upon on this basis.

Alternatively, the steps may be followed in reverse order. In this case, an application type is chosen – for example, an emotion-adaptive music recommendation system – and then the emotions that are considered important for music recommendation are selected, after which the literature is consulted to ascertain which metrics and signals are best suited to recognition of these specific emotions. These steps are reasonably stable across use cases and form the basis of the analytical framework that is introduced

in Chapter 3 of this thesis. Here, we confine ourselves to a brief survey of types of physiological computing applications.

2.3.1 The Four Categories of Physiological Computing

Physiological computing applications can broadly be split into four (conceptual) categories (Cowley et al., 2016), though in practice most physiological computing has been of the classification type (category 1 below) while the fourth category (entrainment) is a somewhat niche area that one might argue is hardly part of physiological computing at all. Below, we offer short descriptions of the categories along with how they are present in the publications that form part of this thesis.

I. Classification

Most physiological computing is based on classifying the affective and cognitive state of the user in line with various physiological signals. For example, a recommender system needs to label, or classify, the affective state of the user during the interaction so that it can recommend items that seem to cause more emotionally positive reactions. For Publication III, we used EEG data to classify words as relevant or irrelevant. In Publication IV, EDA is used to classify text as relevant as well as classify it on the affective scale. For Publication V, we used EDA, EEG, and EKG to assess whether the user found a comic strip funny or not.

II. Prediction

Prediction of the behavior of a user has many potential applications, including detection of whether a driver or pilot is about to fall asleep or experience a medical emergency such as an epileptic seizure. In Publication II, we present predicting users' behavior in an investment situation: will the user bet or not?

III. Biofeedback

Biofeedback has a long tradition in the clinical setting, where it has been used to treat various disorders, both physical and mental. For Publication I, we used biofeedback in combination with a virtual-reality setting to generate a meditation environment that enables a deeper level of meditative experience by means of the user being directly conscious of his or her brain activity, which is fed back as changes in the virtual reality.

- IV. **Entrainment** In entrainment, a user’s physiological state is manipulated by an audio or visual signal toward some desired state, such as relaxation or concentration. Because the entrainment process is rather straightforward and does not demand any computation (for example, when the user simply listens to an audio file), the author would argue that entrainment is more a tool that could be used in physiological computing than a category of it.

Category	Application	Signals	Publications
Classification	Annotation, games	EDA, EEG, EKG	II, IV, VI
Prediction	Games, annotation	EDA, EKG, EEG	II, III, V
Biofeedback	Meditation, games	EEG, EKG	I, VII
Entrainment	-	-	-

Figure 2.4: The categories of physiological computing addressed in the publications.

2.4 Affective Computing

Affective computing is a field that overlaps with physiological computing. Most of what can be considered affective computing is physiological computing, and vice versa.

In her book *Affective Computing*, one of the founders of the field, Rosalind Picard, defines affective computing as “[c]omputing that relates to, arises from, or deliberately influences emotions [...]”. Affective computing includes implementing emotions, and there can aid the development and testing of new and old emotion theories. Affective computing also includes many other things, such as giving a computer the ability to recognize and express emotions, developing its ability to respond intelligently to human emotion, and enabling it to regulate and utilize its emotions” (Picard, 1997). The element that is most relevant to physiological computing is giving computers the ability to recognize human emotions, which can be done in real time by means of physiological sensors. So, in one sense it can be said that physiological computing is an indispensable tool in the arsenal of affective computing. However, it could be said also that affective computing represents just one subset of the logic and application layers of physiological computing: it is one conceptual framework for interpreting

indices such as arousal that come from the index layer below. Affective computing, however, is not simply part of physiological computing; it is possible for affective computing systems to utilize indices and context information not based on physiology.

One major inspiration for affective computing was recognition of how much of our everyday communication is non-verbal while computers are completely unable to comprehend this kind of information. The benefits of being able to measure affective information of the nature would be twofold: the communication bandwidth from human to computer would increase dramatically, and also it would give the computer a kind of affective context information on what the mood of the user is. Whether the computer should offer assistance or, instead, perhaps stay silent might depend greatly on the type of emotions present in the user.

2.4.1 Emotion Theories and Definitions

Emotions, moods, and “affects” are mentioned throughout the thesis, not only in the affective computing section, so it is worth addressing the fact that there is a large amount of ambiguity as to their exact meaning. Indeed, there is, as there has been since ancient times, a large amount of debate and disagreement on their definitions and theories based on them. On the other hand, in common parlance, there is often no disagreement at all, with the terms getting used interchangeably (Batson et al., 1992).

Ekkekakis (2012) divided the affective phenomena into three groups:

- **Core affect:** The primitive, underlying non-reflective feeling often accompanying mood and emotion but not necessarily always available to consciousness. I can also appear alone without mood or emotion. Examples of core affect include pleasure, relaxation, tension, and tiredness. As an example of how these can be utilized in physiological computing, for Publication I relaxation was used as a user state representation during meditation in virtual reality, and the amount of relaxation was connected to the ability to levitate in the virtual world.
- **Emotion:** Emotion is defined as complex inter-related sub-events directed toward a stimulus (in contrast to core affects, which might have no specific target or object). Emotion also needs to generate overt behavior congruent with the emotion (e.g., a smile) and be connected with a cognitive appraisal of the stimulus and its meaning and implications. As one example of how emotions can be used in

physiological computing, for Publication IV the emotional reaction to media content was recorded.

- **Mood:** Mood differs from emotions in that it is usually longer-term and often more global instead of having a specific object as emotions do. Since mood is more of a long-term effect, physiological computing that utilizes moods would have to have a biocybernetic loop that might operate at a slower tempo than usual. One example might be a mindfulness mood journal that gives the individual users feedback on their day-to-day moods.

Sometimes, instead of labeling emotions by category (as, for instance, fear or happiness), it is more useful to think of the affective phenomena as a continuous space. This can be especially useful in quantitative analysis and for some machine learning algorithms that are more suitable for such a continuous and analytic model. The most common such model is the circumplex model of emotion (Larsen and Diener, 1992), which maps the affective phenomena in two-dimensional space of arousal and valence. One benefit of the dimensional model is that it corresponds well with the physiological emotion models and is often used in psychophysiological studies (Lang, 1995).

2.4.2 Decision-Making and Emotions

The traditional view of decision-making posits that emotions and decisions should be kept separate, that emotions can only distort the decision-making process; people have been asked to “think calmly” and “keep a cool head” when making decisions. However, it has been demonstrated that emotions are an essential and mandatory ingredient in the decision-making process: research on patients with brain lesions that disrupt emotions has shown that emotions are an essential part of decision-making. These patients, who had normal levels of intelligence but lacked emotions, were often incapable of even rudimentary decision-making, frequently “getting stuck” without being able to decide which of two options to choose and hence ending up in a kind of endless loop, for they did not even feel boredom that would help them recognize that it was time to quit pondering (Damasio, 1994; Bechara and Damasio, 2005). In more serious decision-making scenarios, such as stock trading, these patients might not be able to learn that a certain investment was bad and could well just keep investing.

To explain this connection between emotions and decision-making, Damasio et al. (1996) proposed the *somatic marker hypothesis* (SMH), which

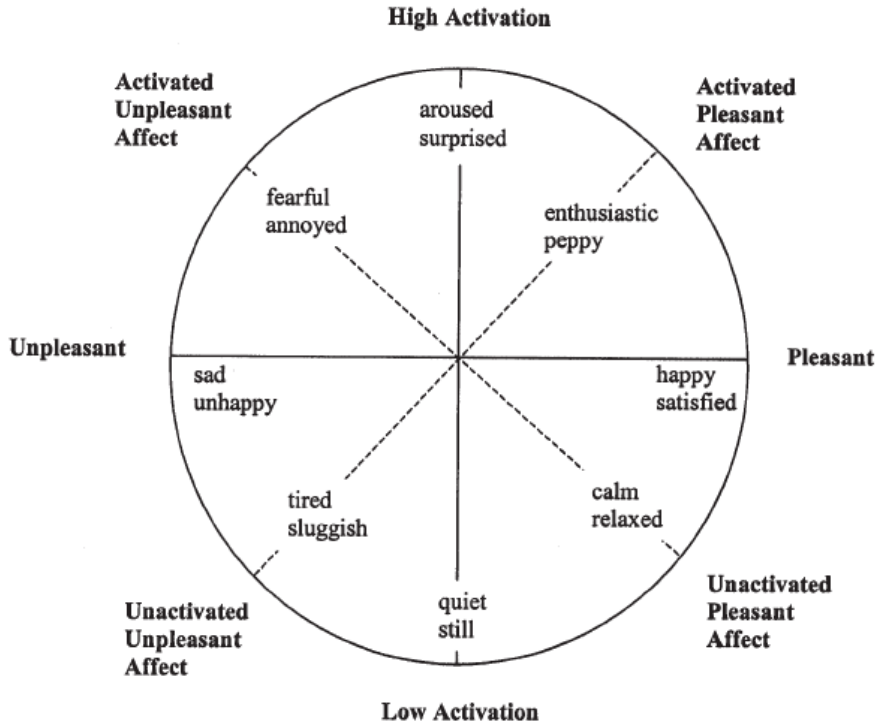


Figure 2.5: Dimensional model of emotions.

postulates that situations and decisions elicit emotions that, in turn, generate bodily, or somatic, responses such as increased heart rate and skin conductivity. The somatic responses can then get associated with these decisions. These somatic markings then act as a kind of heuristic in future decision-making: a decision that has in the past led to negative outcomes becomes marked with a negative somatic response, which leads to avoiding the decision or response. Similarly, situations that have in the past been positively marked can prompt decision-makers to choose actions that lead to these positive outcomes.

If correct, the somatic marker hypothesis leads to interesting possibilities for physiological computing: the somatic markers that guide the decision-making are first expressed as physiological states that are then interpreted by the decision-making process. These physiological states can also be captured by recording the physiological signals of the user. Thus, by observing which somatic marker, or physiological representation of an emotion, is elicited before a decision is made, the system can potentially

predict what type of decision the user will make. For example, if in observing a stock trader a very large positive emotional response is detected, it could be predicted that the broker is going to make an investment decision. In Publication II, behavior of poker players is reported upon in relation to making decisions on whether to bet or not, and it can be seen that physiological responses such as skin conductivity correlated with the player's decisions.

2.5 Wearable Computing

Wearable computing, designing miniature body-borne computational and sensory devices, naturally complements physiological computing. Designing physiological sensors that can be worn as part of everyday clothing instead of being hooked into a non-portable laboratory recording device allows physiological computing to move into the real world. Physiological sensors have already been successfully integrated into “smart clothing,” including armbands (Krause et al., 2003; Lisetti and Nasoz, 2004), shirts (Lee and Chung, 2009), vests (Pandian et al., 2008), and gloves (Ryoo et al., 2005; Peter et al., 2005). Furthermore, it is not always necessary to create new sensors: existing sensors, such as the fitness market's heart-rate bands, can be used in combination with a smartphone or tablet to enable wearable computing applications without additional hardware investments (Healey and Logan, 2005; Oliver and Flores-Mangas, 2006). Also, as the sensory technologies progress, wearable solutions become more and more feasible; for example, wireless versions of EEG have been proven suitable for real-time acquisition and analysis of mental states (Berka et al., 2004).

Wearable computing is especially important for this thesis because of the work's core aim of examining how physiological computing can be utilized across various domains of day-to-day activities. While the empirical studies reported upon in this thesis were done in a laboratory setting, the applications could be reproduced via only wearable devices.

Chapter 3

Research Questions and Method

Physiological computing is a constantly developing and highly varied field of HCI that is still extremely unorganized. To aid in future development and research in physiological computing, a comprehensive view of the whole field is needed that allows transfer of knowledge between projects and research groups. So as to facilitate a broader view, the research questions directed toward exploring physiological computing from three perspectives that complement each other and can together answer basic questions such as when, why, where, and how to use physiological computing.

In this chapter, the rationale for selecting the research method chosen is described. Then the research questions are introduced, along with a five-layer analytical framework that is used to delve into those questions. In the remainder of the thesis, the framework is used to answer the research question as well as show the similarities and differences in physiological computing across application areas and domains.

3.1 Research Methods

The research methodology in this thesis has two levels. Firstly, each of the publications presents a separate case study, with each employing constructive research as outlined in Herbert Simon's *Sciences of the Artificial* (Simon, 1996). The method is also known as *design science* (Peppers et al., 2007). Each of the studies is then further compared to the others by means of a multiple-case-study research design. Below, we will describe the constructive process used in the separate case studies, then address how the multiple-case-study design is used to compare among the individual studies.

3.1.1 Design Science

The design science research process involves six steps (Lehtiranta et al., 2015; Peffers et al., 2007): (1) selecting a problem that is practically relevant, (2) obtaining preliminary understanding of the topic, (3) design and development, (4) testing and demonstration, (5) evaluation, and (6) communication and dissemination of the results. For the sake of brevity, we describe how these steps were implemented in one of the eight case studies, and the details of the other case studies can be found in the accompanying publications. As an example here, we use Publication I, on the neuroadaptive meditation system *RelaWorld*.

Problem Identification

It has been shown that meditation, especially mindfulness meditation, has a wide range of benefits, such as stress reduction (Grossman et al., 2004). However, a problem was identified in that it is often difficult to find suitable space for meditation devoid of visual and auditory distractions, especially during busy office life. The task was, therefore, to explore how virtual reality, as well as wearable physiological computing in the form of neurofeedback, could be used to facilitate novice meditators in their everyday environments.

Obtaining Preliminary Understanding of the Topic

To obtain the initial understanding, a thorough literature review was performed. Because the topic was highly multidisciplinary, the literature review was broken down into specific topics. Firstly, various meditation practices and traditions were explored to find out what would be the most suitable techniques both for virtual reality and for the empirical user study that would be conducted to validate the result. Secondly, the literature on existing technological meditation aids and previous experiments on assisted meditation was surveyed. Thirdly, the literature pertaining to the use of neurofeedback for meditation and also for clinical treatment of relevant conditions such as stress and depression were studied. Finally, the bodies of knowledge obtained from all the separate literature reviews were combined to design the optimal system.

Design and Development

The third step involved the actual construction of the artifact being studied, in this case the virtual-reality neuroadaptive meditation system. The setup

was composed of two main parts. The “back-end” was responsible for recording the EEG signals, processing them into a suitable format, and delivering them to the “front-end” responsible for the virtual reality. The signals had to be pre-processed to remove noise and artifacts, after which the relevant frequency bands were extracted and converted into a stream of two values, reflecting both the relaxation and the concentration of the user. These values were then sent to the Unity3D game engine responsible for implementation of the virtual meditation chamber.

Testing and Demonstration

To test the system, a user study was run with 43 participants who used the neuroadaptive virtual-reality meditation system to perform two distinct meditation exercises.

Evaluation

The system was evaluated in comparison with a control condition wherein the same meditation exercises were performed via a normal computer monitor instead of virtual reality and without any neurofeedback. Performance was measured via two questionnaires, measuring both the success of the meditation itself and *sense of presence*, an index that has been linked to the ability of virtual-reality systems to elicit positive change (Riva et al., 2015). Both questionnaires showed increased performance during the virtual-reality neurofeedback as compared to the control condition.

Communication and Dissemination of the Results

The results and the description of the system design were communicated and disseminated by a publication presented at a high-level conference (see Publication I in this thesis). A similar pattern can be identified in all of the other case studies described in the publications included in the thesis. These case studies will be compared to each other by means of the multiple-case design described in the next section.

3.1.2 Research Design: Case Studies

According to Yin (Yin, 2013), a case study is an

“empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident.”

In this thesis, we examine the contemporary phenomenon of physiological computing as it is utilized in various real-life contexts. Furthermore, the thesis explores boundaries on two levels. Firstly, the biocybernetic loop ties the context and user together such that they are not always trivial to decouple. More importantly, one of the main drivers for this thesis was a desire to study the internal boundaries that exist between the individual levels in the design of physiological computing: the boundary between the raw physiological signals and the features extracted from them, the boundary between the formal metrics and the cognitive/affective constructs that are derived from them, and so on. Indeed, the concept of “boundary,” or “interface,” is central to the sciences of the artificial: an artifact is an interface between the inner and outer environment (Simon, 1996).

The research extends across all levels of artificiality, since the lowest level of the model can be studied as phenomena of natural science: the task is to describe the physiological signals as they are. However, as we move to the higher layers in the analytical model, the concepts studied become increasingly artificial and also the question shifts from what to how: the lower levels involve what IS, while the logic and the application layers deal with “ought.” To paraphrase Simon, “[e]ngineering [...] is concerned not with the necessary but with the contingent – not with how things are but with how they might be – in short, design” (Simon, 1996).

Therefore, the research method has to take into account this broad spectrum of requirements: the need to examine the “real” physiological phenomena of psychophysiological signals, while the design-driven, artificial abstractions of the physiological computing applications build on cognitive/affective constructs that are partly natural and partly artificial. Case studies were chosen as the research method because this approach is suitable when “the theory is weak, occurrences are still scarce and application variations numerous” (Jenkins, 1985). Case studies are also suitable for addressing research questions that start with “how” and “why” (Yin, 2013).

The design of case studies proceeds in three steps: the case is defined; then, the type of the study design is chosen; and, finally, the role of existing theory is considered. In this thesis, the “case” represents an application of physiological computing, and each case contains as embedded sub-cases the various layers of the analytical framework introduced in Section 3.3. The complete design follows a multiple-case design with embedded units as seen in Figure 3.1.

The final step in designing a case study is deciding whether to use existing theories in developing the research questions, selecting cases, and determining which data are relevant (Yin, 2013). This thesis relies heavily

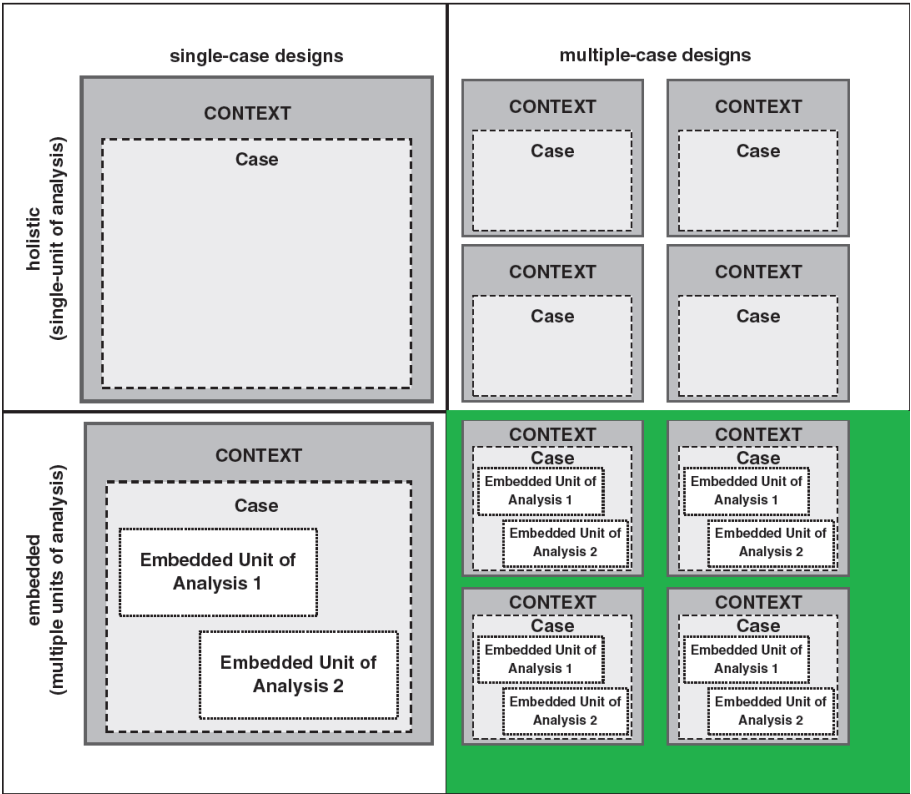


Figure 3.1: Types of case-study designs, from Yin (2013) – the approach for this thesis follows the multiple-case design with embedded sub-units, marked in green in the figure.

on previous work by Pope et al. (1995), Fairclough (2009), and Cowley et al. (2016), alongside others (Novak et al., 2012). The five-layer analytical model that forms the embedded sub-units is built on these existing foundations.

3.2 Research Questions

Work to answer the research questions led us to examine the field of physiological computing from the three key perspectives, which can be formulated via those three questions as presented below.

Motivation for Research Question 1

The first research question was motivated by the realization that the original definition of the biocybernetic loop was not sufficient to describe all or even most use cases of physiological computing: the idea that the physiological signals need to loop back to the user in real time is too restrictive. Out of the eight case studies included in this thesis, only half follow the “immediate feedback to the user” pattern. The “feedback” can be directed instead to the designers of future systems as well as to benefit other users of the system (see Figure 3.2). With the first research question, therefore, we aimed to explore how the concept of the biocybernetic loop can be extended.

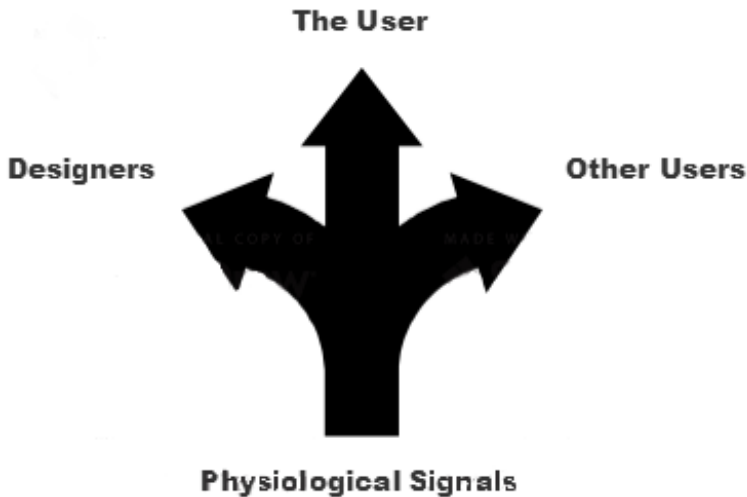


Figure 3.2: Three types of biocybernetic loop.

Research Question 1: How Physiological Computing Can Be Extended beyond the Primitive Biocybernetic Loop

The first research question asks *why* and *when* physiological computing is a suitable approach. Several distinct *roles* that physiological computing can take, such as that of a tool in annotating data for recommender systems and of an aid in technology design, are identified and explored. We also explore various ways the biocybernetic loop can be implemented.

Motivation for Research Question 2

One of the most critical questions in designing physiological computing applications is the decision on how the physiological signals should be interpreted – that is, what kind of representation of user state should be used for mapping the physiological signals. Sometimes a very complicated cognitive/affective construct can be suitable whereas at other times a very simple, even binary, aroused vs. not-aroused representation may suffice. Therefore, the second research question was designed for exploring whether and how the requirements for the user representation vary as the dynamics of the biocybernetic loop change.

Research Question 2: How the Representation of the Psychophysiological State Varies As the Dynamics of the Biocybernetic Loop Change

Physiological computing is based on measuring and adapting to the physiological states of the user. However, there are diverse ways these states can be interpreted from the raw physiological signals. The second research question is intended to address the *what* question by exploring the uses of various psychological and affective constructs that are used to capture and model the user's psychophysiological state. More specifically, with this question we aimed to study how these representations vary when the dynamics of the biocybernetic loop change.

Motivation for Research Question 3

It is often useful to delegate the interpretation of the signals, generation of features, and even the user representation to machine learning algorithms. This can lead to analytically optimal solutions but often also ignore expert knowledge that could be used in deriving the features and deciding on the user representation. Therefore, the third research question was designed for considering whether the differences in the biocybernetic loop affect the feasibility of machine learning approaches.

Research Question 3: How the Dynamics of the Biocybernetic Loop and the Chosen Psychophysiological Representation Affect the Choice of Machine Learning Methods

Physiological computing systems range from simple biofeedback systems, wherein a given physiological signal is mapped directly to some audiovisual cue, to very complicated ones in which several, very different physiological

signals are combined for deriving higher-level affective and cognitive user states that are used further to adapt the system in some way. The third research question addresses the *how* and explores the various approaches that can be taken when one is designing physiological computing systems. Special attention is given to how different machine learning approaches are utilized: sometimes PC systems are based on expert rules derived from the literature such that a given change in a physiological signal is mapped to specific action, while sometimes it is more useful to let the machine learn the best possible way to adapt to the changes in the physiological signals.

3.3 The Five-Layer Model of Physiological Computing

To help us answer the three research questions, an analytical framework was designed to assist in understanding the differences and similarities between the roles, states, and approaches to physiological computing.

With this thesis, we propose a model for physiological computing that builds on the idea of the biocybernetic loop introduced by Pope et al. (1995) and further developed by Fairclough (2009), as well as on the work of Cowley et al. (2016), who derived cognitive and affective indices from well-defined metrics generated from physiological signals. The layer model is a well-known design pattern that allows for horizontal *separation of concerns*, dividing the problem into self-sustained subproblems that can each be solved separately (Goedicke, 1990).

The layers build on top of each other, with the lowest layer dealing with the raw signals and the second layer extracting metrics or features from those signals, which are then converted at the third layer into indices representing cognitive or affective constructs. The fourth layer then combines these outputs into decisions that help the fifth layer to implement some kind of application.

Technically, there is also a sixth layer: the lowest layer contains the actual physical measurement equipment, along with details on how and where each signal should be measured. This layer encompasses such details as what conductive pastes to use with the sensors and so on. These details are beyond the scope of this thesis, and the interested reader can check the work of Cacioppo et al. (2007). In principle, these details could also be considered part of the signal layer.

3.3.1 The Signal Layer

The first layer deals with the physiological signals that have been acquired from the physiological sensors. At this point, the system is dealing with time series of numbers – usually, one real number for each measurement made. Since different signals are often sampled at different rates, the time series can differ between signals in their numbers of values (for example, EDA can be sampled 32 times per second, while EEG and EMG demand sampling rates of 1,000/second or more). Also, some signals, such as the EDA, have only one time-series, while EEG has one series for each electrode, often up to 64 or even more.

One very important part of the signal layer is artifact rejection. For example, the processing of EEG often has several steps, such as removing blinks, which cause artifacts and create noise in the data. Techniques such as the Independent Component Analysis (ICA) have been successfully applied (Hyvärinen et al., 2004) to separate signals corresponding to the actual brain activity from those generated by eye movements. Another common procedure is use of the deconvolution algorithm applied for the EDA signal. Also, some generic cleaning is often necessary, such as applying a band-pass filter to remove the 50 Hz electrical interference caused by all the electrical devices around us that use alternating current, which also pulsates in all the electrical cabling around us. Again, such details are not central to the topic of this thesis, and the interested reader can consult Cacioppo et al. (2007).

The signal layer also deals with topics such as baselining and calibrating the signals. Often when analyzing the signals, we are interested not in absolute measures such as actual skin conductance in Siemens but in the relative change of conductance during an experiment, because the absolute level can change for irrelevant reasons such as room temperature. Hence, at the signal layer the values are centered at some baseline so that they are comparable.

3.3.2 The Metrics Layer

On the next level, the various metrics are extracted from the raw signals. For example, an algorithm parses the raw time series of the EKG signal, picking up the R-spikes that indicate heartbeat, and calculates the heart rate, or, more precisely, the inter-beat interval. Numerous metrics have been developed in psychophysiological research to quantify the activity of the various physiological signals (Cowley et al., 2016). These formulations distill the expert knowledge on the most relevant parts of the signals for

physiological computing. For example, in the complicated EKG signal, which contains several peaks, each with its own shape and amplitude, it has been found that most salient information can be found simply by observing the timing of the R-spikes in order to extract the heart rate. Each signal has such implicit information embedded in the form of the metrics derived from it. In using machine learning methods to automatically generate metrics from a signal, this expert knowledge might not be utilized, but, on the other hand, novel metrics might be found that have not yet been documented.

3.3.3 The Indices Layer

The indices layer deals with the conceptual interpretation of the psychophysiological phenomena: it addresses the question of whether emotions should be interpreted categorically (as fear, anger, etc.) or dimensionally (e.g., valence/arousal). The questions of core affect, mood, and emotion discussed earlier are settled at the indices layer. At this layer, the metrics are used to calculate indices that are useful for the application being developed. For example, the SCR and SCL of EDA can translate to arousal, EEG alpha activity to relaxation, and facial EMG to valence.

While some basic affective states such as “arousal” can be reasonably straightforwardly derived from physiological signals, there is some disagreement as to whether emotions have some universal signature that could be used for detecting them (Kreibig, 2010). Barrett (2006) argued that, since emotions are both context-dependent and constructed, finding *a priori* emotion-specific patterns of autonomic nervous system (ANS) activity is improbable, while meta-analyses by Cacioppo et al. (2007) suggest the possibility of a certain degree of autonomic specificity. Furthermore, some of the indices are inherently context-bound. For example, for publications III and IV, the index being detected is “relevance,” which does not make any sense outside the context. One might say that relevance has intentionality: something in the context is relevant, and talking about relevance without the stimulus that is relevant is meaningless. Therefore, while it might be true generally that emotions do not display a universal signature of the ANS that is always the same for all users, that does not matter as long as the physiological computing system being developed keeps track of the context and claims only that in this given context, given this particular stimulus, the ANS response for this particular emotion will be such and such.

3.3.4 The Logic Layer

The next layer, the logic layer, decides how the indices are going to be used. In some applications, the logic layer might, for example, track the user's level of arousal and trigger some action if the user is getting too excited. Conversely, one of the classic examples of physiological computing is measuring whether pilots or drivers are getting too sleepy to fly/drive and alerting them if necessary. Also, it is possible to combine several indices, possibly for different signals, to generate more intelligent behavior: if we detect high arousal, we also want to know whether the user is showing signs of positive or negative valence, which yield opposite interpretations for the arousal. The logic layer also deals with context-awareness and multimodality – that is, combining input from physiological and non-physiological sources.

Temporal patterns too can be handled at this layer. For example, a hidden Markov model can be built to take into account not only the current physiological state but also the history of such states. The lower levels are primarily concerned with phasic, short-term events, while longer-term tonic phenomena that might take place over a period of minutes can be captured in the logic layer.

More importantly for this thesis, the logic layer is where the differences in the dynamics of the biocybernetic loop occur. Here the information coming from the indices layer is interpreted, and, on the basis of the loop type, either used to adapt the system for the user in real time or forwarded to a recommender system to improve recommendations for all users, as in the case of affective annotation for collaborative filtering. Or perhaps the stimulus–response pair is simply stored for the designers to use as a guide for further development of the system, in which case the adaptation happens at much slower pace.

3.3.5 The Application Layer

The last layer is the application layer, which implements the actual physiological computing application. This final layer implements the four categories of physiological computing: classification, prediction, biofeedback, and entrainment, utilizing the capabilities provided by the logic layer below.

3.3.6 The Full Model

A great deal of confusion exists as to the exact meaning of concepts such as analytical framework, theoretical framework, conceptual framework, theory, and model. The framework described in this thesis shares much in

common with the definition of conceptual framework by Miles and Huberman (1994):

A conceptual framework explains, either graphically or in narrative form, the main things to be studied – the key factors, variables, or constructs – and the presumed interrelationships among them.

However, Ravitch and Riggan (2016) list three other definitions of conceptual framework:

- A conceptual framework can be a visual representation of a study’s organization and major theoretical tenets
- It may be a combination of all the elements of a research process, including the disposition, interest, and positionality of the researcher
- “Conceptual framework” can be just another term for a theoretical framework, in which case the meaning of “conceptual framework” largely depends on what constitutes *theory*

The framework described in this thesis is a bricolage of all these themes. It is intended to create an analytical framework that is helpful both for designing experiments in physiological computing and for serving as a type of “design pattern” for application developers hoping to harness the possibilities of these novel technologies. For the sake of simplicity, the framework shall be simply called the analytical framework here.

Layer	Description
Application	Deals with the implementation details of the actual system, whether it be for classification (annotation), prediction, biofeedback, or entrainment.
Logic	Concerned with topics such as signal fusion (combining input from several physiological signals) as well as context information, user profiling, and multimodality in general. The logic of the biocybernetic loop is contained in this layer and suitable feedback and adaptation specified.
Indices	Uses the metrics to generate indices of the user’s psychophysiological state, such as arousal and concentration. At this layer, the various emotional models and cognitive frameworks are considered.
Metrics	Uses metrics calculated from the raw signals, such as the amplitude of phasic spikes in EDA or the frequency-band powers in EEG.

Signals	Contains low-level details of physiological signals, such as EDA, EEG, and EKG. The layer involves details such as sampling rates and sensor locations (e.g., which EEG sensor locations are suitable for certain purposes).
---------	--

Table 3.1: The Five-Layer Analytical Framework for Physiological Computing.

Chapter 4

Empirical Studies

In this chapter, we describe the application areas where we have employed physiological computing. For each area, we first give a brief description of the area in general, which is followed by presentation of related work showing how others have approached physiological computing in this domain. Then we describe our studies and how they tie in with the research questions, including what physiological signals were used and what indices and metrics were derived. Then we offer our findings. Each of the case studies provides its own part of the answer to the research questions. These partials are combined into a holistic view in Chapter 5.

4.1 Games

Games are an interesting application area to study, for at least two reasons. The first is that the games industry has become one of the biggest forms of entertainment, if not the biggest, in terms of revenue, so it is worth studying games simply to learn how to make new and better ones. The first publication therefore concentrates on pure video gaming and on how physiological computing can be applied to this domain. With that publication, the aim was not to uncover how to use physiological computing in gaming application itself but to consider how physiological computing can be used by game designers to analyze existing game designs and create better games. This publication stands out also in that unsupervised learning was used to analyze the physiological and behavioral data during game play.

Another reason to study games is that much in our everyday activities has game-like aspects and qualities. Finding the fastest route to work and driving as fuel-efficiently as possible is very much a game. Weight-loss programs, diets, and exercise systems too can easily be seen as game-like.

Indeed, there is a whole branch of science that studies “gamification”: how to make everyday activities more like games (Deterding et al., 2011).

Accordingly, the second publication on this topic concentrates on a specific situation that arises both in games and in serious real-life settings: investing. In the most general sense, this is something that happens in our life frequently: we “invest” in some resource with the expectation of some return with some probability. In the publication, we look at the act of investment in the game of poker, and at how a player’s physiology can be used to predict what kind of investment decision will be made (that is, whether the player will bet and how much) given the type of hand (which indicates the player’s belief in how successful the investment is going to be) he or she has.

Poker serves as an interesting test case because it is a combination of playful activity with often very serious investment-related decision-making. In fact, very few other asset-investment scenarios demand such rapid decision-making as online poker. Whereas a professional stock trader might make several investment decisions in minutes, the online poker player may make several decisions within seconds. Indeed, online poker could be described as a micro-level economic decision-making environment that allows assessment of behavior under conditions of risk and uncertainty (Siler, 2010).

In addition to studying physiological behavior during the micro-investment decisions of poker, the experiment involved a biofeedback component to assess whether the decision-making could be supported by real-time feedback on the user’s physiological state. It is a well-known fact that losing emotional control, or going into “tilt,” is a major issue in the game of poker (Browne, 1989). That is, after a big loss, players can go into a state of anxious excitement that leads, through suboptimal decision-making, to further losses. To prevent this kind of chain effect, the players were given feedback on their arousal level through sonified heart rate. The idea was that the audio feedback would alert players when they were too aroused and thereby allow them to regain control. The results of this audio biofeedback work are reported on in Publication VII.

4.1.1 Background and Related Work

Physiological computing has been used to augment games in several ways. One classic example is dynamic difficulty adjustment (DDA), wherein physiological feedback from the user is used to calibrate the difficulty of the game such that it provides the optimal challenge. Ewing et al. (2016) used theta and alpha EEG activity to tune game demandingness, while Liu et al. (2009) used EDA, EMG, and EKG to automatically detect anxiety

and adjust the game difficulty accordingly. The latter resulted in a small improvement in performance and self-reported experience rating relative to a difficulty adjustment system based on player performance. Attempts have been made to combine an even larger number of physiological signals, among them respiration, for better performance of DDA systems (Chanel et al., 2011; Tijs et al., 2008).

Being able to objectively measure the user's emotional state during game play, or in interaction with any entertainment technology, is crucial because what is most important for successful user experience is not the outcome but the process of playing (Pagulayan et al., 2002). That is, often the emotional state of the user is simply a proxy: we wish to optimize the user's cognitive and affective state so that they can perform this or that task better. However, in the case of entertainment, user state is itself the target of the optimization.

Various methods have been utilized in automated emotion detection during game play: Mandryk and Atkins (2007) used fuzzy logic to process EMG, EDA, and HR into valence and arousal, which were then modeled into categories such as fun, challenge, boredom, and frustration. There have also been attempts to predict the type of physiological response a user will display on the basis of interactive events: McQuiggan et al. (2006) utilized decision trees and Bayesian methods to predict EDA and cardiac signals in accordance with the situational context in the interactive environment.

In addition, physiological computing has been used to enhance specific game elements such as the camera vantage points. Yannakakis et al. (2010) derived affective states such as boredom, frustration, excitement, and anxiety from physiological signals (heart rate, blood volume pulse, and skin conductance) to predict the preferences of players for particular camera settings. They further discuss the possibility of real-time adaptive affective camera control based on physiological measurements. Similarly, Yannakakis and Hallam (2008) studied children's play preferences during physical games by recording physiological signals (again, heart rate, blood volume pulse, and skin conductance) to predict their game preferences.

Games can also be used to aid in reaching traditional biofeedback goals. For instance, Bersak et al. (2001) helped players to relax by designing a game where two people compete in a racing video game that measures the relaxation level of the players, with the player who is more able to relax winning. Similar gamification techniques could be combined with other, traditional biofeedback tasks to make them more engaging.

Other possible uses for physiological computing in gaming contexts include automatic adaptation of the game atmosphere. In a horror game, for

instance, the excitement level of the player might be used to maximize the effect of fear inducement. Dekker and Champion (2007) designed a game that was dynamically adapted to players’ physiological state (as measured via EDA) to increase the cinematically augmented “horror” elements: the game adapted its shaders, screen shaking, and creation of new monsters on the basis of the player’s physiological responses.

Games also provide an interesting testing ground for considering the applicability of implicit vs. explicit biofeedback in real-time interaction. Kuikkaniemi et al. (2010) studied how implicitly adapting a first-person shooter game affected the gaming experience relative to a situation wherein players were aware of the biofeedback and were able to consciously attempt to control it. It was found that players preferred the added control of being explicitly aware of the biofeedback process.

4.1.2 The Studies

In this thesis, we explore games from three distinct perspectives. In Publication VI, we look at how physiological computing can be utilized by game developers: by means of unsupervised learning techniques, the physiological responses of players are clustered to show that separation between game events can be based solely on the physiological responses. With publications II and VII, a biofeedback poker experiment is studied in two ways: We explore how physiological signals can be used to predict investment decisions – that is, whether the player bets or not. Secondly, we show that sonified heart-rate biofeedback can be successfully used by players for emotion regulation during game play.

Study 1: Game Design Patterns

Study 1, presented in Publication VI, is part of a larger effort to build a design framework for game designers that includes analysis of player physiology, game patterns, and design principles. What is most relevant for the purposes of this thesis is analysis of physiological signals of players for extraction of patterns of physiological activity that might be relevant for designers. Data were collected from participants playing the game Super Monkey Ball. The physiological signals that were recorded included EDA, fEMG (CS, OO, ZM), and EKG.

Two types of analyses were conducted. Firstly, physiological responses around all important game events were collected in time series of eight seconds, creating a matrix of 5x8 values, one value for each signal at each time point. The data were then clustered via the well-known K-means al-

gorithm, and it was shown that the clusters found in the physiological data corresponded with the most important game events. What makes this analysis particularly interesting is that an unsupervised machine learning algorithm, the K-means algorithm, was used to cluster the unlabeled physiological data.

In the second analysis, the problem of overlapping events was tackled in a novel way. One problem with such a rapidly paced and complicated stimulus complex as video games is that the events that cause physiological responses often occur close in time, so the physiological responses overlap. Separating the overlapping physiological responses is a challenging problem. Researchers have attempted to address it in various ways, with perhaps the most important being the deconvolution approaches in EDA signal analysis, which use mathematical tools to separate among several overlapping phasic spikes of EDA activity (Benedek and Kaernbach, 2010). However, a different approach was tested in this study: instead of trying to separate the individual responses, we clustered the overlapping events into “meta-events,” and the physiological response to the whole cluster of events was considered as one response to one event. An unsupervised learning algorithm, the FP-growth algorithm, was used to cluster events that regularly occurred together as meta-events.

Layer	Description
Application	Extraction of design patterns relevant for game design
Logic	Clustering of frequent patterns in physiological data
Indices	Unsupervised machine learning
Metrics	SCP, CS, ZM, OO, IBI
Signals	EDA, EMG, EKG

Table 4.1: Case Study 1, from Publication VI: Game Design Patterns.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

The study demonstrates a physiological computing approach wherein the physiological feedback from the users is used to provide designers of the system with tools to improve for future systems. Therefore, if we assume that these steps would be taken iteratively, the system describes part of a

more complicated biocybernetic loop wherein the user's responses change the system, which then can be tested again by the users to create a new set of feedback for the designers. The timescale of the adaptation is thus very long. Also, the design loop of collecting physiological data, analyzing the data in order to improve the system, and then collecting further data could involve a different set of users during each iteration of the design loop.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

This study involved an extremely free form of representation of the user state: using unsupervised machine learning, the machine learning algorithm is given complete freedom to construct representative prototypes of physiological responses. In studies wherein supervised machine learning is used, the algorithm is at least given labeled groups of data, but in this case, the algorithm has to work solely from the similarities in the physiological time series.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

The fact that the biocybernetic loop was directed towards the designers allowed selection of the unsupervised machine learning method, because the complete data set was available for the machine learning algorithm and there were no strict running-time requirements for the classification. Furthermore, the fact that the designers were "part of the loop" allowed for a very exploratory approach in which behavioral meta-events were firstly generated and then clustered. That is, the interpretation of the resulting clusters by the designers demands very complex analytical processing. Generation of a fully automatic AI that would implement design changes in accordance with the clustering results is an interesting challenge for future work.

Study 2: Anticipatory EDA Responses

In study 2, reported on in Publication II, the electrodermal responses of participants playing computer poker were recorded. In the experiment, the subjects played 128 hands, with two identical sets of 64 hands. All hands were identical for all subjects. There was a short break, of 3–5 min, after the first set, during which the subjects filled in questionnaires on game experience. A special computer poker program was generated that not only allowed us to design exactly the type of hands dealt to each test participant but also facilitated competition, and therefore motivation, among the test subjects by describing an illusory score board that showed how well the

players were (supposedly) doing when compared to other test participants. Their final performance relative to the others on this list also determined the final score presented to them.

The aim for the study was to see whether the user’s *anticipatory* EDA could be used to predict the bet vs. fold decision. This would be consistent with the somatic marker theory (discussed in Chapter 3), which posits that decision-making is accompanied by *somatic markers* that act as heuristics for the decision process. In this study, we were able to show that arousal of the player before an important decision indeed did correlate with the decision he or she ended up making, as would be expected under the somatic marker hypothesis.

Layer	Description
Application	Basic research: recording and publishing of findings
Logic	Increase in phasic EDA activity interpreted as positive investment decision (betting)
Indices	Arousal
Metrics	SCP
Signals	EDA

Table 4.2: Case Study 2, from Publication II: Anticipatory EDA Response during Poker Play.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

Strictly speaking, the next study involved not physiological computing but a basic research experiment to answer important questions in psychophysiology related to how decision-making can be predicted from physiological signals. However, the study has interesting implications for the biocybernetic loop: the idea behind the traditional biocybernetic loop is that the system *reacts* to the user *after* the user has explicitly or implicitly acted. However, if the somatic marker hypothesis is true and the user action can be predicted, the adaptation can take place before that action is taken. This would allow conversational interfaces wherein the system aims to adapt to not only the current user state but also possible future states. This is an interesting topic for future research.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

The study does not directly address the second research question, but its results do imply that a complete representation of the user might need to include modeling of the possible future states and their predicted probabilities.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

This study used traditional statistical methods for testing the hypothesis that anticipatory EDA predicts decision-making, so it did not directly deal with machine learning methods.

Study 3: Heart-Rate Sonification Biofeedback for Poker

The third study, presented in Publication VII, is connected with the same experiment as study 2. In this study, the effect of biofeedback in the form of sonified heart rate was explored as a way to aid poker players in controlling their arousal. As explained in the previous section, the participants in the experiment played 128 hands, in two identical sets of 64 hands. In one set of 64 hands, the subjects heard their heartbeat sonified, while in the other they heard a simulated heartbeat. The subjects were not aware which condition they were in, and, when briefed after the experiment, they could not identify which of the two conditions was the biofeedback one.

The EDA and facial EMG activity of the players were recorded, and the analysis showed that the players had significantly less EDA and facial activity in the biofeedback condition, as was expected. Furthermore, an exploratory data analysis was conducted to see whether there were individual-to-individual differences between players. While the results of the exploratory analysis have to be taken with a grain of salt, it was found that large differences indeed were clear among participants in the effectiveness of the biofeedback.

First of all, after examination of the raw data, it became evident that there were large differences in overall EDA activity among participants. When this EDA activity was taken into account, via use of the standard deviation of EDA activity per participant as a covariate in the analysis, it emerged that players who had a large amount of EDA activity benefited much more from the biofeedback than those with little EDA activity. In fact, for players showing very little EDA activity, the biofeedback actually increased EDA levels.

Secondly, the efficacy of biofeedback was correlated with the results from the BIS/BAS questionnaire that each participant filled in during the experiment. This questionnaire is divided into two parts, with the BIS portion measuring behavior inhibition in response to impending reward and the BAS portion measuring behavior activation towards a reward. In simpler terms, the BIS covers the tendency to withdraw from a negative situation while the BAS has to do with approaching something perceived as positive. When the BIS/BAS scores were used as covariates in the analysis, it was found that high BIS and high BAS scores each seem to point to greater efficacy of the biofeedback in decreasing arousal.

Layer	Description
Application	Enhanced poker play via enabling players to pay attention to their arousal
Logic	Arousal to sound
Indices	Arousal
Metrics	HR
Signals	EKG

Table 4.3: Case Study 3, from Publication VII: The Poker Study.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

Study 3 represents an example of the traditional biocybernetic loop, with which the system adapts to the user in real time. Accordingly, it does not really address how the biocybernetic loop could be extended. It serves more as a point of comparison for those of the studies that go beyond the normal loop dynamics.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

The traditional single-user loop design allowed for very simple representation of the user as a point in one-dimensional space spanned by arousal. That is, the user is completely represented by a real number indicating his or her level of arousal.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine

learning methods?

In this study, no machine learning was used. The mapping of increase in heart rate to audio biofeedback was straightforward.

4.1.3 Summary

These studies explored several aspects of gaming from a physiological computing perspective. Firstly, it was shown that physiological computing can be successfully used in clustering game play patterns, which could provide game designers with useful information. This is a good example of how the biocybernetic loop can be extended to include system designers. In theory, several design iterations could be carried out, during each of which the physiological responses would be recorded and used to explore the design space. Secondly, the gaming context allowed study of the relationship between emotions and decision-making in a scenario that represented a real-life investment situation. The results point to utility of physiology also for predicting user actions. Thirdly, the use of biofeedback for emotion regulation during game play showed that simply making users mindful of their physiological state can be beneficial, a topic that will be covered in more detail in Section 4.3, which deals with meditation.

4.2 Physiological Annotation and Information Retrieval

Information retrieval (IR) is a domain dealing with search, representation, and manipulation of large collections of electronic text and similar human-language data. Today, IR systems are widespread, with millions of people using them daily to facilitate business, entertainment, and education (Büttcher et al., 2016). The most well-known examples of IR services are web search engines, but information retrieval is an essential part of, for example, digital library systems also. Large companies provide enterprise systems for employers to search corporate data, and desktop search systems allow all users to search personal email messages and files.

4.2.1 Background and Related Work

Manning et al. define information retrieval in their classic text as “finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers)” (Christopher D. Manning and Schütze, 2008). The work presented in this thesis concentrates on satisfying the information need, or,

more precisely, on how to detect from physiological signals alone whether the information need was satisfied by some document of unstructured nature that was retrieved from the large collection. Information retrieval is a broad topic, and comprehensive coverage of it is beyond the scope of this thesis. Therefore, the interested reader is referred to the aforementioned book.

One way to approach this particular type of “satisfaction” is to ask whether the document was “relevant” to the information need of the user. Indeed, a large amount of research has been directed toward utilizing relevance in systems such as “relevance feedback” for enhancing retrieval and “relevance judgment” for annotating things (Saracevic, 2007). Even though relevance is a widely used concept, the underlying question “how does relevance happen?” remains open (Moshfeghi and Jose, 2013).

Annotation is the act of attaching metadata to digital content, which, for instance, allows digital media services to implement more sophisticated content management while also analyzing how the content has been used. For users, it enables enhanced content delivery, personalization, and recommendation systems that can utilize metadata that capture user preferences (Adomavicius and Tuzhilin, 2005; Herlocker et al., 2004). The simplest annotation technique is the fully explicit one: users manually annotate content by bookmarking, rating, tagging, and sharing (Leech, 1997). However, manual feedback can be taxing for the user, who usually takes this approach only when having a clear incentive for doing so. Even then, it has been found, users might not stop to give feedback, even though they know that doing so could benefit them in the future (Kelly and Fu, 2006).

Instead of requiring explicit feedback, one can monitor the users implicitly, which demands no additional user effort; in fact, users are usually unaware that the implicit data are being collected at all. Among the implicit signals are dwell time, clicking, and scrolling behavior (Kelly and Teevan, 2003; Soleymani and Pantic, 2012). Existing applications include YouTube, which bases its recommendation not only on explicit user feedback such as likes or subscriptions but also on keeping track of how long users spent watching any given video (Davidson et al., 2010).

Physiological computing is highly suitable for implicit annotation, since the physiological signals are a rich source of implicit information on how the user perceives any given stimulus. The number annotation types that could be derived from all the various physiological indicators is vast. For our work, we chose to concentrate on two categories that are especially important for work and other day-to-day activities: relevance and affective annotation.

Relevance annotation is designed to determine whether a given stimulus was relevant to the user in the context in which the user was situated at the time. This is obviously very relevant in fields such as information retrieval, and one of the publications deals with the issue of whether it is possible to determine from physiological signals alone whether a given stimulus (in this case, a word) is conceived of as relevant by the user in the context of a scientific search.

Affective annotation, on the other hand, is intended to determine what kind of affective reaction the stimulus or content elicited in the user. In any entertainment domain, it is useful to know how the content affected the user on an emotional level: was the content boring or exciting, and did the user feel happy or sad? Affective annotation can be useful in music recommendation and opinion polling. In another example, the largest Italian newspaper, *Corriere.it*, has been collecting affective feedback for its online news articles. The topics of affective and relevance annotation are not separate, though. It has been shown that adding affective features to a model for detecting relevance increases the relevance predictions' accuracy (Arapakis et al., 2009). Therefore, in Publication IV, we concentrate on the affective annotation of content that has already been labeled as relevant.

Application of physiological computing for annotation is still in its early stages. Also, most work done so far has concentrated on stimuli such as video and audio material (music) that are known to generate strong physiological reactions. However, a large amount of the information consumed by users is still in textual format, whether in newspapers, blogs, or physical books. Text, while less emotionally exciting, is an excellent research topic for just that reason: in analysis of a physiological reaction caused by a video clip, any number of confounding factors might be present. An increase in skin conductance might be prompted simply by something moving rapidly on the screen, which might not have anything to do with the information the research is concerned with. In a textual setting, researchers can usually be certain that the reactions generated are truly responses to the information content alone.

Several attempts have been made to automate the annotation process by using implicit physiological signals. To determine the relevance of videos and documents, Arapakis et al. (2010) analyzed facial expressions captured from video material alone or in combination with EDA and skin temperature (Arapakis et al., 2009). The inclusion of affective features in the models based on implicit relevance metrics increased the accuracy of the relevance predictions.

Moshfeghi and Jose (2013) studied how “traditional” implicit signals such as dwell time could be combined with physiological signals (heart rate, skin temperature, and EEG) and affective features calculated from facial expression captured on video. They were thereby able to gain significant improvements in relevance predictions for a video retrieval engine.

Soleymani et al. (2008) found significant correlations between physiological features (EDA, blood pressure, respiration, and fEMG) and the self-reported emotions of participants watching movie scenes. In addition, Tkalcic et al. (2010) studied how to complement image recommender systems that used generic metadata (i.e., genre, total watching time, etc.) with metadata related to users’ affective state. They found that the addition of affective features improved the system’s performance when compared to use of generic metadata alone.

4.2.2 The Studies

We explored implicit annotation of content from two perspectives. In publications III and IV, we look at how physiological computing can be used to infer relevance of content presented to the user. In Publication III, we pinpoint the actual moment of insight by looking at how single keywords can be analyzed by means of features derived from EEG signals, which enables millisecond-scale temporal resolution for the act of perceiving something to be relevant or not. In the follow-up study, presented in Publication IV, we relaxed the time resolution and sought to determine whether the user found a scientific abstract relevant or not. In the experiment, we also used EDA and facial EMG (corrugator supercilii) instead of EEG to see whether the relevance annotation could be derived also from peripheral physiology.

In the third study, also presented in Publication IV, we went a step further and studied whether given media content, once it was successfully classified as “relevant,” could further be annotated with the affective reaction it generated in the user. In this experiment, users were asked to read news articles and provide explicit feedback after each article about what kind of affective reaction the article generated. The users’ EDA signals were recorded while they were reading the articles.

In the final experiment, presented in Publication V, we delved into a somewhat less popular branch of affective studies and concentrated on detection of *humor appraisal*, the act of finding something funny. Even though humor has been a fundamental part of human civilization since pre-historic times (Polimeni and Reiss, 2006), it has received almost no research in the field of physiological computing.

Study 4: Term-Relevance Prediction from Brain Signals

Because relevance judgments are made in the brain, the most natural and intriguing way of approaching relevance detection is by observing the brain signals directly. Therefore, in study 4, high-precision scientific equipment was used to quantify neural activity across 32 EEG channels from 34 participants. The participants were given a scientific topic, such as “climate change” and then presented with single keywords, one at a time. In this way, the exact temporal moment at which the participants processed each keyword could be assessed without the need for tools such as eye-trackers. For each topic, the users were shown six keywords, which were randomly drawn from a pool of relevant and irrelevant keywords as defined by experts.

The data analysis was based on a large number of features extracted from the EEG signals, which included both frequency-based and event-related features. The features were calculated from a one-second time window after the term was displayed on the screen and baselined on the one-second window just before the term’s presentation. A multi-kernel machine learning approach was used to predict the term relevance from just the physiological data. The classifier was able to improve the relevance prediction by up to 17%.

The changes in EEG were localized to certain areas of the brain, such as Brodmann area 10, which has been associated with a range of cognitive functions from recognition to semantic processing. Furthermore, because the changes were similar across participants, the machine learning system did not require user-specific training or calibration. Hence, it can be used to predict relevance also for previously unseen participants and content.

Layer	Description
Application	Annotation of keywords as relevant or irrelevant
Logic	Classification as relevant or non-relevant in line with machine learning features
Indices	Supervised machine learning
Metrics	Alpha, theta, beta, gamma, ERPs
Signals	EEG

Table 4.4: Case Study 4, from Publication III: Term-Relevance Prediction from Brain Signals.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

This case study used physiological signals for detecting relevance from EEG when users were presented with keywords. The study was aimed at proving detection of relevance to be possible, and the adaptation of the system was left for future research. However, at least two distinct types of biocybernetic loops can already be conceived of that utilize the relevance classification presented. The first is a real-time search engine that employs the user's EEG responses to the presented keywords as a guide for the search. This would be a case of the traditional biocybernetic loop. However, the system could also be used to annotate keywords, in which case the annotations might be utilized by other users, in later search situations.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

A supervised machine learning algorithm was used to discriminate between relevant and irrelevant keywords on the basis of a wide range of EEG features that included both frequency and ERP-based features. The results of this study could technically be used directly in two types of loops: those that go directly back to the user and those that are used for others (as in annotation for collaborative filtering – that is, using the relevance assessments of one user to benefit search or recommender system performance for other users). Thus this study shows that certain solutions for user representation can be utilized with biocybernetic loops with different dynamics.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

As with studies for this thesis that used machine learning, here the set of features was several orders of magnitude larger than in expert-system-type approaches. That is, if observed from the perspective of the five-layer model, the differences are in the metrics layer. The signals and their rudimentary processing (artifact rejection, baselining, etc.) stay the same, but machine learning methods favor a much richer set of features. The larger feature set could lead to improved performance but also makes the system design more complicated. A further factor is that it is very difficult to hand-tailor decision logic for several hundred features, while modern machine learning methods have no difficulty in dealing with them.

Study 5: Exploring Peripheral Physiology As a Predictor of Perceived Relevance in Information Retrieval

With study 2, presented in Publication IV and by Barral et al. (2015), the focus was switched from studying the brain signals directly to measuring peripheral physiology, specifically the EDA and facial EMG (corrugator supercilii or *brow muscle*). These signals were chosen because they are relatively unobtrusive to capture, the equipment needed is only low-cost, and they have been previously linked with psychophysiological functions associated with perceived relevance (Ravaja, 2004; Veldhuizen et al., 2003). In an experiment with 40 participants, these signals were recorded while textual content was being presented during an information retrieval task. The content consisted of the first 40 words of scientific abstracts, which pilot studies showed us were sufficient for participants to decide whether a given abstract is relevant to them or not.

The set of features was generated from an eight-second time window around the time the user gave explicit feedback on whether he or she found the given abstract snippet relevant or not. The time window started two seconds before the feedback and ended six seconds after. This choice of time window was based on the fact that EDA responses can take 1–3 seconds to manifest themselves and up to 6–7 seconds to reach their peak (Dawson et al., 2007). Several features were generated with the aim of capturing the amount of EDA and fEMG activity, as well as its change within this time window. The features were used in a multiple-kernel learning system for prediction with unlabeled abstracts.

The study yielded three main findings:

- I. The time window for predicting relevance from EDA was found to be 4–6 seconds after the explicit relevance judgment was given.
- II. For fEMG (corrugator supercilii), the best time window was from one to two seconds after the relevance judgment.
- III. A classifier predicting results for unseen abstracts was able to generate an improvement of 14% over a random baseline.

Layer	Description
Application	Annotation of scientific abstracts as relevant or irrelevant
Logic	Classification as relevant and non-relevant on the basis of the machine learning features

Indices	Supervised machine learning
Metrics	25 features per signal
Signals	EDA, EMG (corrugator supercilii)

Table 4.5: Case Study 5, from Publication IV: Peripheral Physiology As a Predictor of Perceived Relevance.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

The study showed loop dynamics similar to those seen in case study 4: physiological signals are used for detecting relevance. The only difference is that in this case study the peripheral signals were used instead of the brain activity directly. Therefore, the response to the first research question is the same as that for case study 4.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

Although the setup in case study 5 was very similar to that in study 4, the representation in the former deserves special attention. Instead of using existing, signal-specific metrics, such as the amplitude of a phasic spike as usually used for EDA, the procedure involved very methodically slicing the signals into small features that were defined mathematically. Similar “slicing” could be used for any signal. Therefore, this case study represents in pure form the tendency of machine learning approaches to “throw out” previous expert knowledge and derive new representations that are based wholly on the mathematical properties of the signals.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

The main insight gained from this case study is the same as from case study 4 (as would be expected, since the studies share similar biocybernetic loop dynamics). The largest differences, if one considers things from the perspective of the five-layer analytical model, arise at the metrics level, where the number of signals is much higher than in systems wherein hand-picked metrics are used.

Study 6: Affective Response to a Relevant Text Document

The sixth study moves onward from relevance detection and deals with the question of what kind of affective reaction will be generated by content that presumably is relevant. That is, we further divide the relevance part of the relevance vs. irrelevance dichotomy into several classes of possible affective response (see Figure 4.1). It has been shown that using emotional context can improve the performance of recommender systems (Gonzalez et al., 2007).

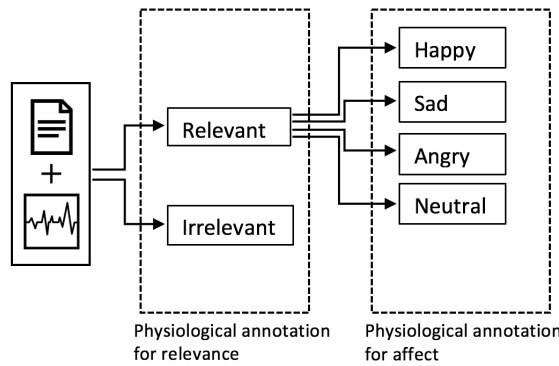


Figure 4.1: Addressing the question of affective relevance annotation, in two steps – firstly, it is determined whether the content is relevant or not (the aim of studies 4 and 5), and then, assuming the content is relevant, the type of affective reaction caused by the content is explored (the aim of studies 6 and 7).

Study 6 was designed to explore the physiological correlates of various affective states that result from reading relevant textual content. The participants ($n=24$) could freely browse a news web site of their choice, select articles they were interested in, and read them for as long as they liked. After reading a news article, the participant was asked to assign the article to one of four affective categories: “happy,” “sad,” “angry,” or “neutral.” The aim with the experiment was to see whether that feedback could be predicted from participants’ physiological data. For this experiment, EDA was the physiological signal of choice, because it has been shown to be indicative of arousal and stimulus novelty (Dawson et al., 2007; Boucsein, 2012).

Since the participants were allowed to spend any amount of time they wished on each item, we had to construct the features in a way that would take into account this variability in trial duration. After down-sampling of

the EDA signal to 32 Hz, the signal was separated into its phasic and tonic components by means of continuous deconvolution analysis as implemented in the Ledalab software (Benedek and Kaernbach, 2010).

For each news article, four features, or metrics, were computed from the EDA signal:

- The overall tonic activity per second, defined as the sum for the tonic component during the time the participant was reading the article divided by the time spent reading
- The amount of phasic activity per second, defined as the sum for the phasic component again divided by the time spent reading
- The average amplitude of the SCRs
- The number of SCRs per second, defined as the total number of SCRs divided by the time spent reading the article

In addition to the physiological data, the participants were profiled by means of the BIS/BAS questionnaire, for examination of the correlation between their personality and the feedback given to the articles. The aim was to explore whether people with differing personality profiles would differ also in the type of affective feedback given, because it has been shown that psychological and personality traits can influence how people access and share content (Celli et al., 2015). The BIS/BAS profiling questionnaire is the same one used in Study VII to profile poker players and their reactivity to audio biofeedback.

Layer	Description
Application	Affective news
Logic	Annotated news
Indices	Emotion categories (happy, sad, angry, neutral)
Metrics	Mean tonic, mean phasic, SCR amplitude, number of SCRs
Signals	EDA

Table 4.6: Case Study 6, from Publication IV: Affective Annotation of News Articles.

Unlike in previous studies wherein a machine learning approach was used to build a predictive model to differentiate between the relevant and irrelevant cases, here a more traditional statistical approach was taken, with the aim being to prove a statistically significant difference in the physiological responses. Because of the highly unbalanced data, we used Linear Mixed Models (LMM) analysis. Mixed models can handle missing observations and unbalanced designs efficiently, resulting in reliability of conclusions that is not possible with more standard statistical measures such as ANOVA (Bagiella et al., 2000).

The results for physiology were somewhat surprising, in that news articles rated as “neutral” and “happy” did not differ significantly in the amount of EDA activity they generated, while the articles denoted as “sad” and “angry,” which could be considered the negative categories, generated *less* EDA than the neutral and positive ones.

The end result is that in the case of affective annotation of news articles, it can be stated that “all arousal is good arousal” since the positively and neutrally rated articles generated more EDA activity than the more negative categories of “angry” and “sad.”

The BIS/BAS questionnaire results showed that the BIS score, which measures inhibition and aversion, correlated negatively with rating articles as “happy.” Meanwhile, BAS scores, measuring appetitive motivations, correlated positively with giving “happy” feedback and negatively with “neutral.” That is, the initial results indicate that personality traits can indeed be used in predicting the type of affective feedback a user will give on news articles. However, since participants were free to choose what articles to read, it could also be that people with certain personality traits *chose* to read articles that were more or less negative.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

This case study offers another example of physiological annotation wherein there is flexibility for the design of the loop dynamics: the annotations could be used to adapt the system to the users themselves or benefit other users in some collaborative scenario. The annotations could also be useful for system designers. For example, in the case of a newspaper that collects affective feedback for its news articles, the feedback could influence the types of content the newspaper produces or perhaps the type of advertisements it attaches to certain type of news.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

The user representation in this study was based on metrics derived from the literature – for instance, related to the amplitude of phasic spikes. Furthermore, knowledge of the underlying mechanism of the EDA signal was used to separate between the tonic and phasic components of that signal and also to separate overlapping phasic spikes. Furthermore, the indices used were emotion categories selected from a “human perspective”: categories such as “happy,” “sad,” and “angry” are concepts arising from human understanding of affective behavior and not derived directly as the mathematically best representatives of the underlying physiological behavior.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

This study did not utilize machine learning at any level of the design. The metrics were derived from the literature, and the affective indices were basic emotions.

Study 7: Physiological Correlates of Humor

The seventh study explored an important aspect of affective annotation that has received surprisingly little attention in the fields of affective and physiological computing, the detection of humor appraisal (again, the act of finding something funny). Humor is a universal human quality: no culture has been found that would be unfamiliar with humor (Polimeni and Reiss, 2006). Humor is not just for fun. It can be crucial in coping with stress and trauma and is considered the most important character trait for satisfaction with life (Kuipera, 2012). Humor has potential to improve performance both in the workplace (Morkes et al., 1999) and in educational settings, where it has been shown to stimulate learning by making it more enjoyable and less stressful (Conkell et al., 1999; Dormann and Biddle, 2006).

In addition to its usefulness in the affective annotation of media content, humor has other uses in HCI. For example, the ability to express and understand humor is crucial for emotional conversational agents (Nijholt, 2007; Khooshabeh et al., 2011; Morkes et al., 1999; Ring et al., 2013). It has even been said that for a computer to pass the Turing test it must develop a sense of humor, so as to understand when the user is serious vs. trying to be funny (Mauldin, 1994).

Previous work has shown that a feeling of amusement is associated with increased activity of the sympathetic nervous system (Martin, 2010), which is reflected in increased heart rate (Foster et al., 1998; Averill, 1969;

Fiacconi and Owen, 2015) and EDA (Foster et al., 1998), both well-known indicators of arousal (Cowley et al., 2016). However, it is obvious that not all states of high arousal are related to humor. Similarly, while humor appraisal is associated with high valence, the opposite is not true: there are positive affective states that have nothing to do with humor. Accordingly, the dimensional model of emotion is not really adequate for capturing the humor phenomenon (Fujimura et al., 2012). Therefore, for this case study we examined humor appraisal in terms of categorical state.

To study the possibility of using physiological computing for humor appraisal, we conducted a user study with 25 participants. The participants browsed humorous web comics while physiological signals (EDA, EEG, and EKG) were being recorded. The aim for the study was twofold. Firstly, we wanted to explore how well physiological signals can be used in the detection of humor appraisal – that is, whether the approach makes any sense. Secondly, on the assumption that the humor appraisal can indeed be detected, we wished to provide a white-box exploration of the underlying physiological responses to humor so that the effect on each signal would be separately explained in a manner allowing future researchers and developers to create such humor detection methods with ease. All too often in physiological computing, a black-box method is used that only provides proof that a certain state, such as humor, can be detected, while giving no indication of precisely how.

To give a comprehensive picture of humor as a phenomenon, we showed how well each signal performs alone as well as how well the machine learning method applied performs when using all the physiological signals together. Furthermore, this was done both by using the participant’s own data to train the system and in a between-user setup wherein other users’ data were used to train the system. Finally, the performance of the system was benchmarked against a state-of-the-art video-based facial emotion recognition system to see how well physiological computing can compete with camera-based approaches.

In this study, we extracted numerous metrics, or features, from the raw signals and then performed feature selection on the training set of the data. The large number of features resulted from both cutting the data in the temporal dimension into windows that represented the start, middle, and post-feedback periods and performing several mathematical transformations that extracted not only the amount of activity in each signal but also the amount of *change* in the activation. This approach yielded some unexpected features that ended up being highly predictive. Specifically, we found that the overall change in gamma-band activity was

highly correlated with humor appraisal. This showed that the approach taken, which could be considered data-mining, was able to extract features that were not addressed in the literature or other previous work.

Layer	Description
Application	Automated annotation of humorous content
Logic	Prediction of whether the user experienced a web comic as funny or not
Indices	Humor appraisal
Metrics	Several hundred features that were derived from the raw signals
Signals	EDA, EEG, EKG

Table 4.7: Case Study 7, from Publication VII: Physiological Correlates of Humor.

The study produced very detailed description of how each signal can be used for humor appraisal. The results can be briefly summarized thus: we found that of the physiological signals EEG performed best, followed closely by EDA, while EKG was not very useful. The combination of all signals performed slightly better than any of the signals alone, and the camera-based system slightly outperformed the full physiological system, but the physiology-based approach provided results that were comparable and proved the feasibility of physiology-based humor appraisal detection in situation wherein camera-based systems would not be applicable (for reasons such as privacy concerns and unavailability of suitable cameras in cases such as mobile use).

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

The case study only provided examples of how humor appraisal can be detected. The actual adaptive system design was left for further work. Ability to detect when the user has found a given stimulus humorous could be very useful for embodied conversational agents, which could thereby tune their sense of humor to be appropriate for a given user. Such an application would be an example of the traditional single-user real-time biocybernetic loop.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

In this study, many features were generated algorithmically, to best capture the information in the underlying signals. Because there was no ready-made construct for the state of “finding something funny,” the task of representing this was left for the machine learning algorithm.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

The case study shows that the choice of machine learning method was not tied very much to the type of biocybernetic loop. It was connected more to the fact that the exact representation of “humor appraisal” and how it might be manifested in the physiological response was less well-known, so a more exploratory approach had to be taken, with the machine learning method being used to generate the representation. This case study supports the idea that there does not exist a clear correlation between the dynamics of the biocybernetic loop and the machine learning methods and, rather, that the machine learning method is more tied to the selected representation of the user state, which, in turn, is not directly tied to the type of the biocybernetic loop.

4.2.3 Summary

The goal with these studies was to explore the insights that physiological computing can yield with regard to users’ information needs and affective responses. The studies produced two important results, which are intertwined: we showed that it is possible to determine with reasonable accuracy whether the textual content was considered to be relevant or not by the user and, if the content was relevant, what kind of affective reaction it caused in the user.

The results have implications for both the field of information retrieval and that of affective recommender systems. Because EDA can be used in predicting relevance both on short, phasic level (as in the case of the scientific abstracts) and on longer, tonic level (as was the case with the news articles), it should be possible in principle to construct an information retrieval system that, perhaps in combination with other physiological signals, annotates content solely on the basis of implicit signals.

4.3 Meditation

In recent years, meditation, especially mindfulness, has become a hot research topic. Even though the exact brain-level mechanisms are still being investigated, the benefits of meditation are multitudinous and well proven. Meditation has been demonstrated to be useful in stress reduction (Chiesa and Serretti, 2009) and management of depression and anger (Grossman et al., 2004), as well as dealing with relationship, attention, and emotionality issues (Sedlmeier et al., 2012). However, irrespective of all these benefits, meditation is not very widely practiced, especially in the Western world. One reason might be that busy modern life and hectic office environments are not exactly conducive to meditation practice. Also, learning to meditate does demand some effort and proper instruction. A virtual-reality system can solve both of these problems.

Neurofeedback has been successfully used in the treatment of various disorders, among them epilepsy (Stermann and Egner, 2006), ADHD (Lubar et al., 1995), learning disabilities (Fernández et al., 2003), and autism spectrum disorders (Kouijzer et al., 2009). While meditation has been used to aid people with similar problems, as in the case of ADHD (Zylowska et al., 2008), little research has been done to explore whether neurofeedback could be combined with meditation techniques.

4.3.1 Background and Related Work

Meditation can be defined as a set of “complex emotional and attentional regulatory strategies developed for various ends, including the cultivation of well-being and emotional balance” (Lutz et al., 2008). Meditation is a broad concept that covers various practices, traditions, and religious-belief systems. Therefore, it is necessary to specify precisely what meditation is being discussed. Many meditations currently practiced, especially in the Western world, draw from Buddhist tradition (Lutz et al., 2008). Of these, the mindfulness-based practices have garnered significant scientific interest and have been proven to be highly successful as well as compatible with the cognitive behavioral framework of psychotherapy (Morgan, 2003).

The differences in meditation techniques exist not only on a theoretical level. They have been shown to elicit different physiological responses in the meditators: meditation techniques aimed at concentration and focused attention have been shown to correlate with beta (13–30 Hz) and gamma (30–50 Hz) EEG activity, while meditation techniques aimed at increasing pure awareness have been linked with activity in the theta (4–8 Hz) and alpha (8–10 Hz) bands (Travis and Shear, 2010).

The traditional view on meditation has been that, to be effective, it has to be engaged in daily and for extended periods of time. However, recent studies in short-term mindfulness meditation programs have shown promising results. Performing twenty minutes of meditation exercises for five consecutive days was already enough to reveal improvements in self-reported scores for depression and anger (Tang et al., 2007). Furthermore, there are indications that even brief mindfulness meditation of three sessions could improve cardiovascular variables such as heart rate (Zeidan et al., 2010) and reduce cardiovascular reactivity to a stressor while the practitioner is performing simple mindfulness meditation exercise (Steffen and Larson, 2015).

A Computer-Assisted Meditation System

Several attempts have already been made at computer-assisted meditation systems. Baños et al. (2012) designed a virtual environment for the elderly that was aimed at generating positive mood. The system consisted of a simulated walk in a virtual park where the users could listen to relaxing melodies, nature sounds, and guided meditation instructions. Self-reporting indicated that the system both reduced negative emotions and increased positive ones.

To help users with no previous meditation experience practice mindfulness meditation, Chittaro and Vianello (2014) designed a mobile application called AEON that allowed the users to enter their thoughts on a mobile device that would then visualize those thoughts as if on parchment submerged in water. Afterward, the users could interact with the touchscreen to produce waves that would slowly wipe out the written thoughts. When the researchers tested the system against normal mindfulness practices, they found that their system was both more pleasant and easier to use than the traditional, non-assisted method. while it was also more effective.

To study the effectiveness of a mobile app for stress reduction, Carissoli et al. (2015) compared the mobile-app users to a control group and a group that listened to relaxing music. While the findings did not reach statistical significance, both the mobile-app users and the group listening to relaxing music did show a trend of decreased stress levels.

Meditation Systems with Virtual Reality and Biofeedback

While the systems mentioned in the section above were proven to be useful, they can be further enhanced by the addition of biofeedback and virtual environments. For example, Meditation Chamber, a virtual environment

for meditation training, used EDA for biofeedback in three guided meditation and relaxation exercises, which utilized a head-mounted display for stereoscopic imagery (Shaw et al., 2010). Similarly, a project called Sensorium built a neurofeedback environment that used the participant's brain waves and heart to produce sounds and control light effects in the room in a way that induced contentment, relaxation, happiness, and inner harmony in the users (Thilo, 2011).

To help chronic-pain patients to cope with their pain, Gromala et al. designed a guided mindfulness walking meditation in a virtual-reality forest where the weather conditions were directly coupled with the user's EDA: relaxing would clear the weather while anxiousness would cause fog to appear (Gromala et al., 2015). Another system, MediAid, used neurofeedback with aural entrainment to aid users in mindfulness practices. The system combined EEG neurofeedback with audio feedback that used binaural beats, a technique that produces two audio signals for headphones such that the signals are slightly out of tune. That is, one ear received an audio tone that was slightly lower-pitched than the signal coming to the other ear. This difference in the signals can be used to entrain the brain oscillations, though the science behind the technique is still maturing (Sas and Chopra, 2015).

4.3.2 The Study

Study 8: RelaWorld, a Neuroadaptive Virtual-Reality Meditation System

The eighth study in the thesis project took us away from the topics of decision-making, investment, and information retrieval, for considering a more holistic view of everyday life: how can physiology help with relaxation, mental health, and wellness and even shed light on topics such as spirituality? For exploration of these topics, a user study with 43 participants was conducted to test the feasibility of combining meditation techniques with modern technologies such as virtual reality and neurofeedback.

The system consisted of a virtual-reality meditation space, specifically a tropical island paradise, that the user could enter simply by wearing the head-mounted display (HMD) and headphones, with the motivation being that the user could escape any hectic work environment for a quick burst of mindfulness in an optimal setting without having access to a dedicated, real-life meditation chamber. In addition to providing a distraction-free environment for meditation practice, the system synthesized findings in the neurofeedback literature to help users obtain optimal mental states

more easily. Because mindfulness meditation is based on the idea of “being mindful” of bodily and mental states, the neurofeedback that allowed users to pay attention to even deeper aspects of their psyche as measured by EEG felt like a natural addition to the mindfulness practice.

Layer	Description
Application	Neuroadaptive virtual-reality meditation
Logic	Relaxation & concentration to movement in VR
Indices	Relaxation & concentration
Metrics	Alpha, theta
Signals	EEG

Table 4.8: Case Study 8, from Publication I: RelaWorld.

Study 8 followed a within-subject design wherein each participant used three types of meditation system: the full neuroadaptive virtual-reality setup, virtual reality without neuroadaptation, and a control condition wherein participants saw the meditation environment on a monitor screen (without neuroadaptation). The experiment setup used further division into two types of meditation, which were performed in each of the conditions, but, since the meditation exercises did not display any differences, they were combined in the data analysis. The system was assessed by means of two questionnaires: a subset of the ITC-Sense of Presence Inventory (Lessiter et al., 2001) and the Meditation Depth Questionnaire, MEDEQ (Piron, 2001).

The meditation depth questionnaire revealed a general trend: the full system with VR and neurofeedback was ranked as best, and the VR system without neurofeedback was rated lower than the full system but higher than the control condition of monitor screen without neurofeedback. Similar results were obtained with the Sense of Presence questionnaire: The combination of VR and neurofeedback produced the greatest sense of presence, followed by the virtual reality without neurofeedback. The control condition of a computer monitor without neurofeedback elicited the lowest level of presence.

How this study addresses the research questions:

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

The study represents a classical case of single-user real-time adaptive biofeedback.

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

The user representation is built on expert knowledge: the user state is represented as a point in two-dimensional space spanned by concentration and relaxation, and these two indices are directly inferred from EEG frequencies in accordance with findings from previous research.

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

This case study is an example of a traditional biofeedback application wherein user state is used to adapt the system in real time to the user. Furthermore, the representation uses a full expert-system approach wherein the metrics and indices are based on outputs from previous research and preset manually. Accordingly, no machine learning was used in this case study.

4.3.3 Summary

This part of the thesis project was devoted to examining how physiological computing can be used to transform what would be a boring biofeedback application into an immersive virtual-reality application. Helping users to relax by providing biofeedback is a classic task Critchley et al. (2001), but the normal biofeedback design involves only a simple visual cue to inform the user of his or her progress. This study showed how physiological computing can be used to combine the biofeedback task with much more meaningful interaction.

The study also showed how this type of passive BCI (Brain Computer Interface) can be used without training or calibration, in that the system used well-known EEG features to perform the biofeedback.

Chapter 5

Findings

In the previous chapter, eight empirical studies were presented that explored the possibilities of physiological computing from several, quite different angles. Each of the case studies represented an attempt to give a partial answer to the research questions presented in Chapter 3. The aim with this chapter, in turn, is to bring the partial answers together and thereby provide a holistic view of physiological computing and the various dynamics of the biocybernetic loop.

5.1 Purpose: Extending the Concept of the Biocybernetic Loop

Research Question 1: How can physiological computing be extended beyond the primitive biocybernetic loop?

With the first research question, we explored how the dynamics of the biocybernetic loop can be extended beyond the case wherein the physiological computing system adapts to the user’s own signals in real time, which is the way it is traditionally described (Pope et al., 1995; Fairclough, 2009). It is possible also for the physiological responses of the user to be directed to the system designers to aid in future development, or for those responses to be used to adapt the system for all users, as when physiological signals are used in annotation for collaborative filtering.

Only two of the eight case studies (studies 3 and 8) strictly fit the “classical definition” of direct, real-time adaptation for the user, while in the other six cases either the loop was not “directed” back to the user or the adaptation did not occur instantly.

One way to approach the question is to use a conceptual framework for interactive systems as proposed by Solovey et al. (2015). In this framework,

interactive systems are categorized by their functional level and immediacy. At functional level, interactive systems are divided into those that adapt *syntactically* and those that adapt *semantically*; that is, the system can either adapt the user interface itself (syntactic) or adjust the content presented through that interface (semantic). The second way to group interactive systems is by their two levels of immediacy: the system can either adapt the currently shown interface and content or perform the adaptation at some point in the future. In Table 5.1, the case studies are listed alongside the type of biocybernetic loop they implemented. Also presented are the functional and temporal categories.

Study	Loop type	Functional	Temporal
1) Game design patterns	Designers	Syntactic	Future
2) Anticipatory EDA responses	Designers	Syntactic	Future
3) Biofeedback for poker	User	Semantic	Immediate
4) Relevance from brain signals	User/others	Semantic	Future
5) Peripheral relevance	User/others	Semantic	Future
6) Affective text annotation	User/others	Semantic	Future
7) Humor detection	User/others	Semantic	Future
8) RelaWorld	User	Semantic	Immediate

Table 5.1: The Types of Biocybernetic Loops in the Case Studies, with Their Functional and Temporal Categorizations.

The first thing to notice is that those cases following the traditional biofeedback approach and implementing the real-time, single-user biocybernetic loop (again, studies 3 and 8) are both of the type *semantic* and *immediate*. Indeed, of all the case studies, only the “traditional” ones fall into the temporally immediate category. Another point to note is that only

the *designer* loop types have a syntactic adaptation, while the rest of the case studies fall into the semantic category. It can also be seen that the temporally immediate adaptations are always semantic, and the functionally syntactic adaptations always occur in the future, with the only overlap being in the semantic adaptation happening both immediately and in the future. Furthermore, the temporal class of the semantic adaptation depends on the dynamics of the biocybernetic loop – those loops that do not return directly to the user lie in the future category.

5.2 Representation: How User State Is Modeled

Research Question 2: How does the representation of the psychophysiological state vary when the dynamics of the biocybernetic loop change?

To explore Research Question 2 by using the analytical model introduced in Chapter 3, we can look at the bottom three layers of the model, which deal with how physiological signals are mapped to cognitive and affective concepts. See Table 5.2.

Study	Signal	Metrics	Indices	Loop type
1) Game design patterns	EDA, fEMG, EKG	Auto-generated feature	Game patterns	Designers, future, syntactic
2) Anticipatory EDA responses	EDA	Phasic EDA	Decision	Designers, future, both
3) Biofeedback for poker	EDA	SCP	Arousal	User, now, semantic
4) Relevance from brain signals	EEG	Auto-generated feature	Relevance	Both, future, semantic
5) Peripheral relevance	EDA, fEMG	Auto-generated features	Relevance	Both, future, semantic
6) Affective text annotation	EDA	Auto-generated feature	Affective state	Both, future, semantic
7) Humor detection	EDA, EKG, EEG	Auto-generated feature	Humor appraisal	Both, future, semantic

8) RelaWorld	EEG	Alpha, theta	Relaxation, concentra- tion	User, now, semantic
--------------	-----	-----------------	-----------------------------------	------------------------

Table 5.2: The Representation of the User State with Various Loop Types.

Firstly, we can see that the signal layer is not affected by the overall design of the system. For example, EDA was used in all types of biocybernetic loops and user representations; that is, the case studies did not show preference for certain signals for specific types of biocybernetic loops or user representations.

The changes between systems start to emerge when the metrics are extracted from the signals. The traditional biofeedback setups, such as those in studies 3 and 8, use manually specified metrics that are based on previous research, such as studies of alpha- and theta-band activity of EEG and the phasic spikes of EDA. However, in those case studies wherein the adaptation was not immediate, the metric selection was automated to generate features that extract maximal information from the signals. Also, the number of metrics was much higher in these cases: it is easy for an algorithm to generate hundreds of features and then use feature selection methods to pick the ones that seem most informative. Furthermore, the machine learning methods can learn decision boundaries that utilize a large number of features.

The only case in which pre-defined metrics were used in a non-immediate, non-user loop was case study 2, wherein we studied the anticipatory EDA responses to decision-making. Here, the knowledge from previous work was so strong that the experiment was a confirmatory one rather than exploratory. Furthermore, the signal studied, EDA, is considerably simpler than, for example, EEG.

One important distinction in terms of the user representation is whether the indices are *context-bound* or *unbound*. Some indices, such as arousal, relaxation, and concentration, are not explicitly bound to the context and environment, while others, such as *relevance* and *humor appraisal*, are tightly coupled with the context: it makes sense to represent the user as being “relaxed,” but it is not sensible to have a representation of “a user in a state of finding something relevant.” In phenomenological terms, the bound indices have *intentionality*: they are always about something.

From a practical point of view, a system can be designed that detects whether the user is relaxed or aroused without any need to know what kind of stimulus made that user relaxed. This has several practical consequences. For example, an unbound system can be designed without the designers having to worry about issues such as synchronization, which can be an extremely tricky issue, especially when one is dealing with wireless sensors and highly time-specific physiological features such as the event-related potentials in EEG studies.

5.3 Approaches: When and How to Use Machine Learning

Research Question 3: How do the dynamics of the biocybernetic loop and the chosen psychophysiological representation affect the choice of machine learning methods?

To articulate the answer to the third research question, we can again look at how the choice of loop dynamics affects the various layers. One thing to notice is that the signal layer is not affected by whatever loop type is chosen. However, when extracting metrics from the signals, we can see a pattern emerge. In the “traditional” biocybernetic loops that adapt to the user in real time, the calculated features are pre-defined and few in number: in the study wherein biofeedback was used to aid poker players, the only feature extracted from the EKG signal was heart rate, which was then directly sonified. In the RelaWorld study, the EEG signal from six electrodes was compressed into two numbers, one from theta activity and the other from alpha. In contrast, in the cases in which the loop was not real-time adaptive, the feature generation was much more thorough, often with hundreds of features being generated that were then fed into machine learning algorithms.

However, whether the adaptation was syntactic or semantic or whether the loop was directed to the designers, other users, or even the users themselves had no effect on the metrics’ generation. The aim was to extract from the signals as much information as possible, and this process was, in essence, always the same.

Study	Representation	Loop type	Machine learning approach
1) Game design patterns	Algorithmically generated clusters (categorical)	Designers, future, syntactic	Unsupervised

2) Anticipatory EDA responses	Literature-based (dimensional)	Designers, future, both	Expert
3) Biofeedback for poker	Literature-based (dimensional)	User, now, semantic	Expert
4) Relevance from brain signals	Algorithmically generated features (categorical)	Both, future, semantic	Supervised
5) Peripheral relevance	Algorithmically generated features (categorical)	Both, future, semantic	Supervised
6) Affective text annotation	Algorithmically generated features (categorical)	Both, future, semantic	Supervised
7) Humor detection	Algorithmically generated features (categorical)	Both, future, semantic	Supervised
8) RelaWorld	Literature-based (dimensional)	User, now, semantic	Expert

Table 5.3: The Various Machine Learning Approaches Used in the Case Studies.

Looking at Table 5.3, we can see that in these cases the representation of the user was strongly dependent on whether machine learning was used. In the expert systems, the features were also pre-defined and based on previous research. For example, in the RelaWorld study, the relaxation index was tied to the alpha-frequency-band activity in line with a literature review. However, in each of the cases wherein machine learning was used, it was already present at the metrics layer: the machine learning method was utilized to generate optimal features instead of building on top of pre-defined features.

Whether the user representation was categorical (as with basic emotions such as “happy” and “angry” in case study 6) or dimensional (as in the case of the relaxation and concentration in study 8) showed a pattern. The dimensional model was used in the expert systems, whereas categorical classification was used in the machine learning approaches.

Chapter 6

Discussion

Physiological computing is an emergent field with potential to both improve existing human–computer interaction designs and provide wholly new paradigms such as affective computing. However, the field is still unorganized, and there is a notable lack of standards and transfer of knowledge between researchers and developers. The aim behind this thesis was to provide an overview of the field from several perspectives and to explore what the differences and similarities are between various approaches to physiological computing. More specifically, we examined the details and dynamics of the biocybernetic loop that has been shown to be the fundamental basis of physiological computing (Pope et al., 1995; Fairclough, 2009).

This dissertation contributes to HCI in general and physiological computing in particular by extending the scope of physiological computing beyond the single-user real-time biocybernetic loop, to encompass systems wherein the feedback might not be imminent and the loop might not go directly to the users themselves (for instance, the physiological signals might be used to assist in the design of new systems and annotate data for other users).

For better understanding of the various aspects of physiological computing, we have presented a five-layer analytical framework, which we used in answering the three research questions for the thesis. That is, through it, we dealt, firstly, with the different dynamics of the biocybernetic loop; secondly, with how these differences are reflected in how the user is represented; and, finally, with how machine learning methods can be applied with different biocybernetic loop and representation types.

6.1 Summary of the Main Findings

The main findings are best explored in terms of the research questions and the associated five-layer analytical model. The first research question deals with how physiological computing can be extended beyond the basic concept of the single-user biocybernetic loop that adapts the system in real time. The main finding in this connection is that in many cases the physiological computing system does not actually adapt in real time; the adaptation can happen when, for example, the physiological responses of the users are used by the designers of the system to improve that system. In this case, the loop “points to” the designers of the system instead of the user, and the timescale can be weeks instead of milliseconds. The physiological responses can also be used in a collaborative manner – for instance, when used to annotate data for recommender systems. Thus, the first main finding is that the concept of physiological computing, and the biocybernetic loop, can be extended to deal with a wider array of application domains.

The second research question deals with the various ways the physiological signals can be interpreted and used in building representations of the user’s cognitive or affective states, along with, more precisely, how these representations change when the dynamics of the biocybernetic loop, as explored in line with Research Question 1, change. Traditionally, there have been at least two main ways of handling user state representation: dimensional and categorical. In a dimensional paradigm, the user is mapped into a space spanned by some cognitive indices such as valence and arousal. The user representation is then a continuous value. In the categorical paradigm, on the other hand, the user state is classified as belonging to a categorical state – for example, specifying that the user is “happy” or “angry.” Apart from these “classical” approaches, there is the possibility also of leaving the interpretation of the user state to a machine learning algorithm and considering the user a black box. For example, when building a classifier to detect whether a given content is relevant to the user, the machine learning algorithm can detect the best possible features for indicating that the user judged a content relevant instead of first trying to deduce how aroused the user was, or whether users were “happy” or “sad.” The second main finding is that the eight case studies in the thesis project seem to indicate that the traditional dimensional and categorical presentations prevail when the biocybernetic loop too is traditional (i.e., when the loop returns to the user in real time). One tentative explanation for this is that real-time algorithms are easier to build around simpler representations, whereas more complicated machine learning and data-mining techniques are suitable when, for

instance, one is using physiological computing to inform designers about how to improve the product.

Finally, the third research question deals with an issue raised in the previous paragraph: how the choice to use machine learning methods differs between different biocybernetic loops and user representations. Here it is useful to look at the five-layer analytical model to see where exactly the differences occur between these cases. Comparison between the case studies show that the most important differences arise in feature selection and generation: the “traditional” real-time biocybernetic loops use a pre-defined, easily calculable set of features that are derived from existing literature in psychophysiology, such as heart-rate and phasic spikes in skin conductance response. In contrast, the methods based on machine learning and data-mining often use rigorous mathematical tools to generate hundreds of features with the aim of capturing all the information present in the signals, then use statistical feature selection methods to choose the most appropriate ones.

6.1.1 Implications of the Research

In line with earlier research in the physiological computing field, my findings show that the biocybernetic loop as an abstract construct is a sensible basis for modeling physiological computing. However, the findings enable extending the prior notion of the loop into a broader range of application scenarios wherein the adaptation might not be instantaneous or directly affect the user. The analytical framework used in answering the research question provides insights into the differences between biocybernetic loops with different dynamics, and it can also be used as a design heuristic to decide what kind of approach to take when, for example, selecting signals and features for a certain physiological computing application or experiment.

The case studies show that *time* is a very important consideration when one is designing physiological computing, in at least two ways. Firstly, whether the system needs to adapt in real time affects the types of machine learning methods that are suitable. Some machine learning methods can be too computationally expensive to be used in real-time adaptation, especially in wearable computing situations, in which computational resources can be scarce. Furthermore, many machine learning methods, and physiological phenomena, demand training data from the user before they can be used. Hence, the user must either train the system initially or wait until it has gathered enough training samples during actual use.

6.1.2 Limitations

The experiments presented in this thesis were conducted mostly with high-precision scientific equipment. Today there is a wide range of consumer-grade wearable measurement devices on the market that, while enabling the deployment of physiological computing applications in the real world, might also introduce additional challenges – the consumer devices might yield lower signal quality. It would, therefore, be beneficial to see how well the results presented in the publications generalize when taken out of the lab, both as far as recording devices go and with regard to the possible increase in the number of confounding factors when users are allowed to use a system freely wherever and whenever they choose.

The multiple-case-study approach was biased in the sense that it involved sampling data only from a pool of studies by one researcher – namely, the author. A further meta-review should be conducted to explore whether the findings discussed in this dissertation generalize across the whole field of physiological computing.

While the thesis represents an effort to provide details on when machine learning could be useful in designing physiological computing systems, it has not gone into detail as to what specific machine learning methods would be useful in a particular situation.

6.1.3 Directions for the Future

Instead of just looking at abstract models, we should also look at concrete ways to apply these abstractions. Physiological computing needs a central repository of standards, best practices, example data, and code. My dream would be to build an actual, working web-based resource that would contain code, data, negative results, etc. One inspiration could be the Requests for Comments and Best Current Practices from the networking field.

An initial attempt at implementing the web-based repository for physiological computing has been started in efforts to provide an interface based on the five-layer model presented in this thesis. The interface would allow users to easily browse the existing work on physiological computing on the basis of parameters such as signals, metrics, and specific user representations. For example, the user might choose to list only studies that used “arousal” that was detected from “EDA.” The next step for future work would be to carry out a comprehensive meta-review of existing studies in physiological computing, for ascertaining how well they fit the model described in this thesis, and to input the information into the new repository.

The use of machine learning in a broad range of physiological computing scenarios demands a comprehensive research effort for mapping which machine learning methods are suitable for real-time adaptation, which can be pre-trained and which need to be trained with data from the current user. Furthermore, these studies should be done in a transparent way that shows the specific importance of the signal behavior and features. Too often in physiological computing, the machine learning is performed in a black-box manner that shows only that certain signals can be used in, for example, classifying a certain affective state, but not how exactly the signals changed between these states. With Publication VII, we attempted to provide an example of how to describe in detail what signals are useful (in the humor detection scenario), while also comparing the suitability of pre-training from other users to that of using the data from the current user. However, similar work needs to be done in all domains of physiological computing, and in a systematic manner.

References

- Abowd, G. D., Mynatt, E. D., and Rodden, T. (2002). The human experience [of ubiquitous computing]. *IEEE pervasive computing*, 1(1):48–57.
- Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6):734–749.
- Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1.
- Andreassi, J. L. (2000). Psychophysiology: Human behavior and physiological response. 4th edition. *Psychology Press*.
- Anttonen, J. and Surakka, V. (2005). Emotions and heart rate while sitting on a chair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 491–499, New York, NY, USA. ACM.
- Arapakis, I., Athanasakos, K., and Jose, J. M. (2010). A comparison of general vs personalised affective models for the prediction of topical relevance. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pages 371–378, New York, NY, USA. ACM.
- Arapakis, I., Konstantas, I., and Jose, J. M. (2009). Using facial expressions and peripheral physiological signals as implicit indicators of topical relevance. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM 09, pages 461–470. ACM.
- Averill, J. R. (1969). Autonomic Response Patterns During Sadness and Mirth. *Psychophysiology*, 5(4):399–414.

- Baños, R. M., Etchemendy, E., Castilla, D., García-Palacios, A., Quero, S., and Botella, C. (2012). Positive mood induction procedures for virtual environments designed for elderly people. *Interact. Comput.*, 24(3):131–138.
- Bagiella, E., Sloan, R. P., and Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37:13–20.
- Barral, O., Eugster, M. J., Ruotsalo, T., Spapé, M. M., Kosunen, I., Ravaja, N., Kaski, S., and Jacucci, G. (2015). Exploring peripheral physiology as a predictor of perceived relevance in information retrieval. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pages 389–399. ACM.
- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on psychological science*, 1(1):28–58.
- Batson, C., Shaw, L., and Oleson, K. (1992). Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions. *Review of personality and social psychology*, 13:294–326.
- Bechara, A. and Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52(2):336 – 372. Special Issue on Neuroeconomics.
- Benedek, M. and Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658.
- Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., and Olmstead, R. (2004). Real-time analysis of eeg indexes of alertness, cognition, and memory acquired with a wireless eeg headset. *International Journal of Human-Computer Interaction*, 17(2):151–170.
- Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonald, B., and Karkun, R. (2001). Intelligent biofeedback using an immersive competitive environment. Paper at the Designing Ubiquitous Computing Games Workshop at UbiComp.
- Boucsein, W. (2012). *Electrodermal Activity*. Springer US, Boston, MA.
- Bousefsaf, F., Maaoui, C., and Pruski, A. (2014). Remote detection of mental workload changes using cardiac parameters assessed with a low-cost webcam. *Computers in biology and medicine*, 53:154–163.

- Boyer, M., Cummings, M., Spence, L. B., and Solovey, E. T. (2015). Investigating mental workload changes in a long duration supervisory control task. *Interacting with Computers*, 27(5):512–520.
- Bradley, M. M. and Lang, P. J. (2000). Emotion and motivation. *Handbook of psychophysiology*, 2:602–642.
- Browne, B. R. (1989). Going on tilt: Frequent poker players and control. *Journal of gambling behavior*, 5(1):3–21.
- Büttcher, S., Clarke, C. L., and Cormack, G. V. (2016). *Information retrieval: Implementing and evaluating search engines*. Mit Press.
- Cacioppo, J. T., Tassinary, L. G., and Berntson, G., editors (2007). *Handbook of Psychophysiology*. Cambridge University Press, third edition. Cambridge Books Online.
- Carissoli, C., Villani, D., and Riva, G. (2015). Does a meditation protocol supported by a mobile application help people reduce stress? suggestions from a controlled pragmatic trial. *Cyberpsychology, Behavior, and Social Networking*.
- Celli, F., Ghosh, A., Alam, F., and Riccardi, G. (2015). In the mood for sharing contents: Emotions, personality and interaction styles in the diffusion of news. *Information Processing & Management*, pages –.
- Chanel, G., Rebetez, C., Bétrancourt, M., and Pun, T. (2011). Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41(6):1052–1063.
- Chiesa, A. and Serretti, A. (2009). Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. *The journal of alternative and complementary medicine*, 15(5):593–600.
- Chittaro, L. and Vianello, A. (2014). Computer-supported mindfulness: Evaluation of a mobile thought distancing application on naive meditators. *International Journal of Human-Computer Studies*, 72(3):337 – 348.
- Christopher D. Manning, P. R. and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Conkell, C., Imwold, C., and Ratliffe, T. (1999). The effects of humor on communicating fitness concepts to high school students. *Physical Educator*, 56(1):8.

- Cowley, B., Filetti, M., Lukander, K., Torniainen, J., Henelius, A., Ahonen, L., Barral, O., Kosunen, I., Valtonen, T., Huotilainen, M., Ravaja, N., and Jacucci, G. (2016). The psychophysiology primer: A guide to methods and a broad review with a focus on human-computer interaction. *Foundations and Trends Human-Computer Interaction*, 9(3-4):151–308.
- Critchley, H. D., Melmed, R. N., Featherstone, E., Mathias, C. J., and Dolan, R. J. (2001). Brain activity during biofeedback relaxation: a functional neuroimaging investigation. *Brain*, 124(5):1003–1012.
- da Silva, F. L. (2013). {EEG} and meg: Relevance to neuroscience. *Neuron*, 80(5):1112 – 1128.
- Damasio, A. (1994). *Descartes’ error*. London: Vintage.
- Damasio, A. R., Everitt, B. J., and Bishop, D. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 351(1346):1413–1420.
- Davidson, J., Liebal, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., and Sampath, D. (2010). The youtube video recommendation system. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, RecSys ’10, pages 293–296, New York, NY, USA. ACM.
- Dawson, M. E., Schell, A. M., and Filion, D. L. (2007). The electrodermal system. *Handbook of psychophysiology*, 2:200–223.
- Dekker, A. and Champion, E. (2007). Please biofeed the zombies: Enhancing the gameplay and display of a horror game using biofeedback. In *DiGRA Conference*.
- Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM.
- Dormann, C. and Biddle, R. (2006). Humour in game-based learning. *Learning, Media and Technology*, 31(4):411–424.
- Edmonds, E., Everitt, D., Macaulay, M., and Turner, G. (2004). On physiological computing with an application in interactive art. *Interacting with computers*, 16(5):897–915.

- Ekkekakis, P. (2012). The measurement of affect, mood, and emotion in exercise psychology. In *G. Tenenbaum, R.C. Eklund, and A. Kamata (Eds.), Measurement in sport and exercise psychology*, pages 321–332.
- Eugster, M. J., Ruotsalo, T., Spapé, M. M., Kosunen, I., Barral, O., Ravaja, N., Jacucci, G., and Kaski, S. (2014). Predicting term-relevance from brain signals. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 425–434. ACM.
- Ewing, K. C., Fairclough, S. H., and Gilleade, K. (2016). Evaluation of an adaptive game that uses eeg measures validated during the design process as inputs to a biocybernetic loop. *Frontiers in human neuroscience*, 10.
- Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers*, 21(1-2):133–145.
- Fernández, T., Herrera, W., Harmony, T., Díaz-Comas, L., Santiago, E., Sánchez, L., Bosch, J., Fernández-Bouzas, A., Otero, G., Ricardo-Garcell, J., Barraza, C., Aubert, E., Galán, L., and Valdés, P. (2003). Eeg and behavioral changes following neurofeedback treatment in learning disabled children. *Clinical EEG and Neuroscience*, 34(3):145–152.
- Fiacconi, C. M. and Owen, A. M. (2015). Using psychophysiological measures to examine the temporal profile of verbal humor elicitation. *PLOS ONE*, 10(9):1–16.
- Foster, P. S., Smith, E. W. L., and Webster, D. G. (1998). The psychophysiological differentiation of actual, imagined, and recollected mirth. *Imagination, Cognition and Personality*, 18(3):189–203.
- Fujimura, T., Sato, W., and Okanoya, K. (2012). Subcategories of positive emotion. *PSYCHOLOGIA*, 55(1):1–8.
- Gamelin, F. X., Berthoin, S., and Bosquet, L. (2006). Validity of the polar s810 heart rate monitor to measure rr intervals at rest. *Medicine & Science in Sports & Exercise*, 38(5):887–893.
- Goedicke, M. (1990). Paradigms of modular system development. In *Managing Complexity in Software Engineering*, pages 1–20.
- Gonzalez, G., De La Rosa, J. L., Montaner, M., and Delfin, S. (2007). Embedding emotional context in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, pages 845–852. IEEE.

- Gromala, D., Tong, X., Choo, A., Karamnejad, M., and Shaw, C. D. (2015). The virtual meditative walk: Virtual reality therapy for chronic pain management. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 521–524, New York, NY, USA. ACM.
- Grossman, P., Niemann, L., Schmidt, S., and Walach, H. (2004). Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of Psychosomatic Research*, 57(1):35 – 43.
- Hancock, P. A., Pepe, A. A., and Murphy, L. L. (2005). Hedonomics: The power of positive and pleasurable ergonomics. *Ergonomics in Design*, 13(1):8–14.
- Healey, J. and Logan, B. (2005). Wearable wellness monitoring using ecg and accelerometer data. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, pages 220–221. IEEE.
- Herlocker, J. L., Konstan, J. A., Terveen, L. G., and Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22(1):5–53.
- Hettinger, L. J., Branco, P., Encarnacao, L. M., and Bonato, P. (2003). Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. *Theoretical Issues in Ergonomics Science*, 4(1-2):220–237.
- Hyvärinen, A., Karhunen, J., and Oja, E. (2004). *Independent component analysis*, volume 46. John Wiley & Sons.
- Jacucci, G., Spagnolli, A., Freeman, J., and Gamberini, L. (2014). *Symbiotic Interaction: A Critical Definition and Comparison to other Human-Computer Paradigms*, pages 3–20. Springer International Publishing, Cham.
- Jenkins, A. M. (1985). Research methodologies and mis research. *Research methods in information systems*, pages 103–117.
- Kaipainen, M., Ravaja, N., Tikka, P., Vuori, R., Pugliese, R., Rapino, M., and Takala, T. (2011). Enactive systems and enactive media: embodied human-machine coupling beyond interfaces. *Leonardo*, 44(5):433–438.
- Kelly, D. and Fu, X. (2006). Elicitation of term relevance feedback: An investigation of term source and context. In *Proceedings of the 29th Annual*

- International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, pages 453–460, New York, NY, USA. ACM.
- Kelly, D. and Teevan, J. (2003). Implicit feedback for inferring user preference: A bibliography. *SIGIR Forum*, 37(2):18–28.
- Khooshabeh, P., McCall, C., Gandhe, S., Gratch, J., and Blascovich, J. (2011). Does it matter if a computer jokes. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 77–86, New York, NY, USA. ACM.
- Kouijzer, M. E., de Moor, J. M., Gerrits, B. J., Buitelaar, J. K., and van Schie, H. T. (2009). Long-term effects of neurofeedback treatment in autism. *Research in Autism Spectrum Disorders*, 3(2):496 – 501.
- Krause, A., Siewiorek, D. P., Smailagic, A., and Farringdon, J. (2003). Un-supervised, dynamic identification of physiological and activity context in wearable computing. In *ISWC*, volume 3, page 88.
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3):394–421.
- Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I., and Ravaja, N. (2010). The influence of implicit and explicit biofeedback in first-person shooter games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 859–868. ACM.
- Kuipera, N. A. (2012). Humor and resiliency: Towards a process model of coping and growth. *Europe's Journal of Psychology*, 8(3):475–491.
- Lal, S. K. and Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, 55(3):173 – 194.
- Lang, P. J. (1995). Emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5):372–385.
- Larsen, R. R. and Diener, E. (1992). Promises and problems with the circumplex model of emotion. *Review of personality and social psychology*, pages 25–59.
- Lee, Y.-D. and Chung, W.-Y. (2009). Wireless sensor network based wearable smart shirt for ubiquitous health and activity monitoring. *Sensors and Actuators B: Chemical*, 140(2):390–395.

- Leech, G. (1997). Introducing corpus annotation. In Garside, R., Leech, G., and Mcenery, A., editors, *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Addison Wesley Longman.
- Lehtiranta, L., Junnonen, J.-M., Kärnä, S., and Pekuri, L. (2015). The constructive research approach: Problem solving for complex projects. *Designs, methods and practices for research of project management*, pages 95–106.
- Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence: Teleoper. Virtual Environ.*, 10(3):282–297.
- Lisetti, C. L. and Nasoz, F. (2004). Using noninvasive wearable computers to recognize human emotions from physiological signals. *Eurasip Journal on Applied Signal Processing*, 2004(11):1672–1687.
- Liu, C., Agrawal, P., Sarkar, N., and Chen, S. (2009). Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction*, 25(6):506–529.
- Lubar, J., Swartwood, M., Swartwood, J., and O’Donnell, P. (1995). Evaluation of the effectiveness of eeg neurofeedback training for adhd in a clinical setting as measured by changes in t.o.v.a. scores, behavioral ratings, and wisc-r performance. *Biofeedback and Self-regulation*, 20(1):83–99.
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, 2nd edition*. MIT Press.
- Lutz, A., Slagter, H. A., Dunne, J. D., and Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in Cognitive Science*, 12(4):163–169.
- Maguire, M. (2001). Methods to support human-centred design. *International journal of human-computer studies*, 55(4):587–634.
- Mandryk, R. L. and Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International journal of human-computer studies*, 65(4):329–347.
- Martin, R. A. (2010). *The psychology of humor: An integrative approach*. Academic press.

- Mauldin, M. L. (1994). Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition. In *AAAI*, volume 94, pages 16–21.
- McQuiggan, S. W., Lee, S., and Lester, J. C. (2006). Predicting user physiological response for interactive environments: An inductive approach. *AIIDE*, 2006:60–65.
- Miles, M. B. and Huberman, A. M. (1994). Qualitative data analysis: A sourcebook. *Beverly Hills: Sage Publications*.
- Morgan, D. (2003). Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse.
- Morkes, J., Kernal, H. K., and Nass, C. (1999). Effects of Humor in Task-Oriented Human-Computer Interaction and Computer-Mediated Communication: A Direct Test of SRCT Theory. *Human-Computer Interaction*, 14(4):395–435.
- Moshfeghi, Y. and Jose, J. M. (2013). An effective implicit relevance feedback technique using affective, physiological and behavioural features. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13, pages 133–142, New York, NY, USA. ACM.
- Nacke, L. E., Kalyn, M., Lough, C., and Mandryk, R. L. (2011). Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 103–112. ACM.
- Nijholt, A. (2007). Conversational Agents and the Construction of Humorous Acts. *Conversational Informatics: An Engineering Approach*, pages 19–47.
- Novak, D., Mihelj, M., and Munih, M. (2012). A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting with computers*, 24(3):154–172.
- Oliver, N. and Flores-Mangas, F. (2006). Healthgear: a real-time wearable system for monitoring and analyzing physiological signals. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE.
- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., and Fuller, T. (2002). *User-centered design in games*. CRC Press Boca Raton, FL.

- Pandian, P., Mohanavelu, K., Safeer, K., Kotresh, T., Shakunthala, D., Gopal, P., and Padaki, V. (2008). Smart vest: Wearable multi-parameter remote physiological monitoring system. *Medical engineering & physics*, 30(4):466–477.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3):45–77.
- Peter, C., Ebert, E., and Beikirch, H. (2005). A wearable multi-sensor system for mobile acquisition of emotion-related physiological data. In *International Conference on Affective Computing and Intelligent Interaction*, pages 691–698. Springer.
- Picard, R. (1997). *Affective computing*. MIT Press, Cambridge, MA.
- Piron, H. (2001). The meditation depth index (medi) and the meditation depth questionnaire (medeq). *Journal for Meditation and Meditation Research*, 1(1):69–92.
- Polimeni, J. and Reiss, J. P. (2006). The first joke: Exploring the evolutionary origins of humor. *Evolutionary Psychology*, 4(1):147470490600400129.
- Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, 40(1-2):187–195.
- Ramos, C., Augusto, J. C., and Shapiro, D. (2008). Ambient intelligence - the next step for artificial intelligence. *IEEE Intelligent Systems*, 23(2):15–18.
- Ravaja, N. (2004). Contributions of Psychophysiology to Media Research: Review and Recommendations. *Media Psychology*, 6(2):193–235.
- Ravitch, S. M. and Riggan, M. (2016). *Reason & rigor: How conceptual frameworks guide research*. Sage Publications.
- Ring, L., Barry, B., Totzke, K., and Bickmore, T. (2013). Addressing loneliness and isolation in older adults: Proactive affective agents provide better support. *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, pages 61–66.
- Riva, G., Botella, C., Baños, R., Mantovani, F., García-Palacios, A., Quero, S., Serino, S., Triberti, S., Repetto, C., Dakanalis, A., et al. (2015).

- Presence-inducing media for mental health applications. In *Immersed in media*, pages 283–332. Springer.
- Ruotsalo, T., Jacucci, G., Myllymäki, P., and Kaski, S. (2014). Interactive intent modeling: Information discovery beyond search. *Communications of the ACM*, 58(1):86–92.
- Ryoo, D.-W., Kim, Y.-S., and Lee, J.-W. (2005). Wearable systems for service based on physiological signals. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 2437–2440. IEEE.
- Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. part iii: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):2126–2144.
- Sas, C. and Chopra, R. (2015). Meditaid: a wearable adaptive neurofeedback-based system for training mindfulness state. *Personal and Ubiquitous Computing*, 19(7):1169–1182.
- Satyanarayanan, M. (2001). Pervasive computing: Vision and challenges. *IEEE Personal communications*, 8(4):10–17.
- Sedlmeier, P., Eberth, J., Schwarz, M., Zimmermann, D., Haarig, F., Jaeger, S., and Kunze, S. (2012). The psychological effects of meditation: A meta-analysis. *Psychological Bulletin*, 138(6):1139–1171.
- Shaw, C., Gromala, D., and Song, M. (2010). The meditation chamber: Towards self-modulation. *Metaplasticity in Virtual Worlds: Aesthetics and Semantics Concepts*, pages 121–133.
- Siler, K. (2010). Social and psychological challenges of poker. *Journal of Gambling Studies*, 26(3):401–420.
- Simon, H. A. (1996). *The sciences of the artificial*. MIT press.
- Soleymani, M., Chanel, G., Kierkels, J., and Pun, T. (2008). Affective characterization of movie scenes based on multimedia content analysis and user’s physiological emotional responses. In *Multimedia, 2008. ISM 2008. Tenth IEEE International Symposium on*, pages 228–235.
- Soleymani, M. and Pantic, M. (2012). Human-centered implicit tagging: Overview and perspectives. In *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pages 3304–3309.

- Solovey, E., Afergan, D., Peck, E. M., Hincks, S. W., and Jacob, R. J. K. (2015). Designing implicit interfaces for physiological computing: Guidelines and lessons learned using fnirs. *ACM Trans. Comput.-Hum. Interact.*, 21(6):35:1–35:27.
- Solovey, E. T., Zec, M., Garcia Perez, E. A., Reimer, B., and Mehler, B. (2014). Classifying driver workload using physiological and driving performance data: Two field studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 4057–4066. ACM.
- Spapé, M. M., Filetti, M., Eugster, M. J., Jacucci, G., and Ravaja, N. (2015a). Human computer interaction meets psychophysiology: A critical perspective. In *Symbiotic Interaction*, volume 9359 of *Lecture Notes in Computer Science*, pages 145–158. Springer.
- Spapé, M. M., Hoggan, E. E., Jacucci, G., and Ravaja, N. (2015b). The meaning of the virtual midas touch: An erp study in economic decision making. *Psychophysiology*, 52(3):378–387.
- Spapé, M. M., Kivikangas, J. M., Järvelä, S., Kosunen, I., Jacucci, G., and Ravaja, N. (2013). Keep your opponents close: Social context affects eeg and fmg linkage in a turn-based computer game. *PLoS ONE*, 8(11).
- Steffen, P. R. and Larson, M. J. (2015). A brief mindfulness exercise reduces cardiovascular reactivity during a laboratory stressor paradigm. *Mindfulness*, 6(4):803–811.
- Sterman, M. and Egner, T. (2006). Foundation and practice of neurofeedback for the treatment of epilepsy. *Applied Psychophysiology and Biofeedback*, 31(1):21–35.
- Tang, Y. Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q., Sui, D., Rothbart, M. K., Fan, M., and Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences*, 104.
- Tassinary, L. G. and Cacioppo, J. T. (2000). The skeletomotor system: surface electromyography. In Cacioppo, J. T., Tassinary, L. G., and Berntson, G. G., editors, *Handbook of Psychophysiology*, pages 163–199. Cambridge University Press, New York, NY, 2nd edition.
- Thilo, H. (2011). The sensorium: A multimodal neurofeedback environment. *Advances in Human-Computer Interaction*.

- Tijs, T., Brokken, D., and Ijsselsteijn, W. (2008). Dynamic game balancing by recognizing affect. *Fun and games*, pages 88–93.
- Tkalcic, M., Burnik, U., and KoÅir, A. (2010). Using affective parameters in a content-based recommender system for images. *User Modeling and User-Adapted Interaction*, 20(4):279–311.
- Travis, F. and Shear, J. (2010). Focused attention, open monitoring and automatic self-transcending: Categories to organize meditations from vedic, buddhist and chinese traditions. *Consciousness and cognition*, 19(4):1110–1118.
- Veldhuizen, I., Gaillard, A., and De Vries, J. (2003). The influence of mental fatigue on facial emg activity during a simulated workday. *Biological Psychology*, 63(1):59–78.
- Yannakakis, G. N. and Hallam, J. (2008). Entertainment modeling through physiology in physical play. *International Journal of Human-Computer Studies*, 66(10):741–755.
- Yannakakis, G. N., Martínez, H. P., and Jhala, A. (2010). Towards affective camera control in games. *User Modeling and User-Adapted Interaction*, 20(4):313–340.
- Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications.
- Zeidan, F., Johnson, S. K., Gordon, N. S., and Goolkasian, P. (2010). Effects of brief and sham mindfulness meditation on mood and cardiovascular variables. *The Journal of Alternative and Complementary Medicine*, 16:867–873.
- Zylowska, L., Ackerman, D. L., Yang, M. H., Futrell, J. L., Horton, N. L., Hale, T. S., Pataki, C., and Smalley, S. L. (2008). Mindfulness meditation training in adults and adolescents with adhd: a feasibility study. *Journal of Attention Disorders*, 11(6):737–746.